

Fr m: Li, Ruixiang
Sent: Wednesday, March 19, 2003 10:56 AM
T : STIC-Biotech/ChemLib
Subject: Sequence search of Application NO: 09/727,739

Please do a standard search on:

- (i) SEQ ID NOS: 15 and 18 against the interference amino acid databases;
(ii) SEQ ID NOS: 17 and 19 against both commercial and interference amino acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10D19
306-0282

RECEIVED
MAR 19 2003
TECH/CHEM. DIVISION
(STIC)

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 3/25/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 21.4378 Seconds
(without alignments)
276.816 Million cell updates/sec

Title: US-09-727-739B-15

Perfect score: 597

Sequence: 1 MRVSQTHCALLLGLAIC.....PPRRKAGCKNEYWGFSTC 111

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCTR_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTRUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160.5	26.9	110	10	US-09-766-396-3
2	160.5	26.9	110	12	US-10-062-375-3
3	105.5	17.7	140	10	US-09-280-030-64
4	88	14.7	200	9	US-10-101-487-53
5	87	14.6	112	10	US-09-766-396-2
6	87	14.6	112	12	US-10-062-375-2
7	83	13.9	85	10	US-09-766-396-6
8	83	13.9	85	12	US-10-062-375-6
9	80.5	13.5	105	10	US-09-766-396-26
10	80.5	13.5	105	12	US-10-062-375-26
11	80.5	13.5	155	9	US-10-066-500-132
12	80.5	13.5	155	9	US-10-028-072-380
13	80.5	13.5	155	9	US-10-121-049-380
14	80.5	13.5	155	9	US-10-123-904-380
15	80.5	13.5	155	9	US-10-140-470-380
16	80.5	13.5	155	9	US-10-175-746-380
17	80.5	13.5	155	9	US-10-176-918-380
18	80.5	13.5	155	9	US-10-176-921-380
19	80.5	13.5	155	9	US-10-002-796-132

20	80.5	13.5	155	9	US-10-066-273-132	Sequence 132, App
21	80.5	13.5	155	9	US-10-066-494-132	Sequence 132, App
22	80.5	13.5	155	9	US-10-137-865-380	Sequence 380, App
23	80.5	13.5	155	9	US-10-140-474-380	Sequence 380, App
24	80.5	13.5	155	9	US-10-142-431-380	Sequence 380, App
25	80.5	13.5	155	9	US-10-143-114-380	Sequence 380, App
26	80.5	13.5	155	9	US-10-066-269-132	Sequence 132, App
27	80.5	13.5	155	9	US-10-140-002-380	Sequence 380, App
28	80.5	13.5	155	9	US-10-066-193-132	Sequence 132, App
29	80.5	13.5	155	9	US-10-066-211-132	Sequence 380, App
30	80.5	13.5	155	9	US-10-142-419-380	Sequence 380, App
31	80.5	13.5	155	9	US-10-123-262-380	Sequence 380, App
32	80.5	13.5	155	9	US-10-142-423-380	Sequence 380, App
33	79.5	13.3	109	10	US-09-766-396-5	Sequence 5, Appli
34	79.5	13.3	109	12	US-10-062-375-5	Sequence 5, Appli
35	78	13.1	15	9	US-10-101-487-52	Sequence 52, Appli
36	78	13.1	29	10	US-09-766-396-7	Sequence 7, Appli
37	78	13.1	29	10	US-09-766-396-11	Sequence 11, Appli
38	78	13.1	29	12	US-10-062-375-7	Sequence 7, Appli
39	78	13.1	29	12	US-10-062-375-11	Sequence 11, Appli
40	78	13.1	84	10	US-09-766-396-10	Sequence 10, Appli
41	78	13.1	84	12	US-10-062-375-10	Sequence 10, Appli
42	77	12.9	14	9	US-10-101-487-37	Sequence 37, Appli
43	77	12.9	14	9	US-09-316-505-2	Sequence 2, Appli
44	67.5	11.3	95	10	US-09-864-761-41361	Sequence 41361, A
45	67	11.2	133	9	US-10-114-893-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1.

US-09-766-396-3
; Sequence 3, Application US/09766396
; Patent No. US20020013456A1

GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.

Siggnis, George R.
Henriksen, Steven J.

TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 No. US20020013456A1 Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/09/766,396

FILING DATE: 18-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/857,389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-766-396-3

Query Match 26.9%; Score 160.5; DB 10; Length 110;
Best Local Similarity 38.3%; Pred. No. 7.9e-11;
Matches 46; Conservative 17; Mismatches 30; Indels 27; Gaps 6;

QY 8 CALALLGLAALICS-OGAASOPDLASRLQLRALAALPHRSGVSEWRWTFYFNCPL 66
DB 2 CALAALCIVLALGGVGTGAPSDPRL---RQFLQKSLAA---TGKQELAKYFLAE---L 50
QY 67 RWRPRKVKGPOLKAKE-----DLERSVDNLP---PRKAGCKNFYKGFSTSC 111
DB 51 LSEPNQTDALPELDPOAAEQDEMRLQRSANSNPAMPAPRERKAGCKNFWKFTTSC 110

RESULT 2
US-10-062-375-3
Sequence 3 Application US/10062375
Patent No. US20020133000A1
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
de Lecea, Luis
Siggins, George R.
Henriksen, Steven J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 No. US20020133000A1th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/062,375
FILING DATE: 30-Jan-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/857,389
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schmonsees, William
REGISTRATION NUMBER: 31,796
REFERENCE/DOCKET NUMBER: 22908-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-7041
TELEFAX: (415) 324-0638
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-062-375-3

Query Match 26.9%; Score 160.5; DB 12; Length 110;
Best Local Similarity 38.3%; Pred. No. 7.9e-11;
Matches 46; Conservative 17; Mismatches 30; Indels 27; Gaps 6;

QY 8 CALALLGLAALICS-OGAASOPDLASRLQLRALAALPHRSGVSEWRWTFYFNCPL 66
DB 51 LSEPNQTDALPELDPOAAEQDEMRLQRSANSNPAMPAPRERKAGCKNFWKFTTSC 110

DB 2 CALAALCIVLALGGVGTGAPSDPRL---RQFLQKSLAA---TGKQELAKYFLAE---L 50
QY 67 RWRPRKVKGPOLKAKE-----DLERSVDNLP---PRKAGCKNFYKGFSTSC 111
DB 51 LSEPNQTDALPELDPOAAEQDEMRLQRSANSNPAMPAPRERKAGCKNFWKFTTSC 110

RESULT 3
US-09-280-030-64
Sequence 64 Application US/09280030A
Patent No. US20010021515A1
GENERAL INFORMATION:
APPLICANT: Sato, Seiji
APPLICANT: Higashikuni, Naohiko
APPLICANT: Kudo, Toshiyuki
APPLICANT: Kondo, Masaki
TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
TITLE OF INVENTION: DNAS
FILE REFERENCE: 382.1026
CURRENT APPLICATION NUMBER: US/09/280,030A
CURRENT FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: JP10-87339/1998
EARLIER FILING DATE: 1998-03-31
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 64
LENGTH: 140
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Designated is
OTHER INFORMATION: an amino acid sequence of
OTHER INFORMATION: MWPSP-MWPMP20-(His)6-Egf-TEV-Somatostatin 28
US-09-280-030-64

Query Match 17.7%; Score 105.5; DB 10; Length 140;
Best Local Similarity 28.0%; Pred. No. 0.00015;
Matches 42; Conservative 18; Mismatches 35; Indels 55; Gaps 8;

QY 4 SQHICALALIGLALICSGGAA--SQPDLDASRLQLRALAALPHRSGVSEWRWTFYP 61
DB 4 SVLASALALTVAPMAFAAEARTTAKMDMDMEKTVHHH-----HHNSDSE----- 51
QY 62 NCP-----CL-----RWRPRKVKGPOLK---AKEDLE 85
DB 52 -CPLSHDGYCLHDGVCWYIEALDKYACNVGVIGERCQYRLKWWELRDYDPTTENLY 110
QY 86 -RSVDNLP---PRKAGCKNFYKGFSTSC 111
DB 111 FQSANSNPAMPAPRERKAGCKNFWKFTTSC 140

RESULT 4
US-10-101-487-53
Sequence 53 Application US/10101487
Patent No. US20020169125A1
GENERAL INFORMATION:
APPLICANT: LEUNG, DAVID W.
APPLICANT: BERGMAN, PHILIP A.
APPLICANT: LOFOUST, ALAN
APPLICANT: PIETZ, GREGORY E.
APPLICANT: TOMPKINS, CHRISTOPHER K.
APPLICANT: WAGGONER, JR., DAVID W.
TITLE OF INVENTION: RECOMBINANT PRODUCTION OF POLYANIONIC POLYMERS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 077319/0329
CURRENT APPLICATION NUMBER: US/10/101,487
CURRENT FILING DATE: 2002-03-20
PRIOR APPLICATION NUMBER: 60/277,705
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 53
LENGTH: 200
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic fusion
OTHER INFORMATION: protein
US-10-101-487-53

Query Match 14.7%; Score 88; DB 9; Length 200;
Best Local Similarity 48.4%; Pred. No. 0.021;
Matches 15; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 81 KEDLERSVDNLPERRKAGCKNFYKGTSC 111
DB 170 EEEEEEEEEEEEEAGCKNFYKGTSC 200

RESULT 5

US-09-766-396-2

Sequence 2, Application US/09766396

Patent No. US20020013456A1

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

de Lecea, Luis

Siggin, George R.

Henriksen, Steven J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/766.396

FILING DATE: 18-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/857,389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 112 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-766-396-2

Query Match 14.6%; Score 87; DB 10; Length 112;

Best Local Similarity 25.2%; Pred. No. 0.014;

Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSOGAASQPDLDASRLRLQALAAALPHRSVSR

DB 4 CSTRGRKPSALSLLLLLSGIAASALPLESGPTGDSVODATGGRRTGLTLFLAWH-- 61

QY 60 YPNCPCPLRWPRKVG-----POLKAKEDLERSVDNLPERRKAGCKNFYKGTSC 111
DB 62 -----EWASQSSSTAPEGTPELSKRQ--ERPPLQPPHDKKPKCNFFWKTFSSC 111

RESULT 6

US-10-062-375-2

Sequence 2, Application US/10062375

Patent No. US20020133000A1

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

de Lecea, Luis

Siggin, George R.

Henriksen, Steven J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No. US20020133000A1th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/062.375

FILING DATE: 30-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/857.389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 112 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-062-375-2

Query Match 14.6%; Score 87; DB 12; Length 112;

Best Local Similarity 25.2%; Pred. No. 0.014;

Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSOGAASQPDLDASRLRLQALAAALPHRSVSR-----WRTF 59

DB 4 CSTRGRKPSALSLLLLLSGIAASALPLESGPTGDSVODATGGRRTGLTLFLAWH-- 61

QY 60 YPNCPCPLRWPRKVG-----POLKAKEDLERSVDNLPERRKAGCKNFYKGTSC 111

DB 62 -----EWASQSSSTAPEGTPELSKRQ--ERPPLQPPHDKKPKCNFFWKTFSSC 111

RESULT 7

US-09-766-396-6

Sequence 6, Application US/09766396

Patent No. US20020013456A1

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

de Lecea, Luis

Siggin, George R.

LENGTH: 105 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-766-396-26

Query Match 13.5%; Score 80.5; DB 10; Length 105;
Best Local Similarity 27.7%; Pred. No. 0.07;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQALAAALPHRSVGSERWTFYPCPLR-----WRPKVKGPOL-KAK 81
DB 11 LLGATATAALPLEGGTGRDSEHMQEAAGIRKSSLLTFLAWFWFNTSQASAGPLIGEEA 70
QY 82 EDLRSVDNLP-----RERKAGCKNFYKGTSC 111
DB 71 REVARREGAPQOSARRDRMPCRNFFWKTSSC 104

RESULT 10

US-10-062-375-26
Sequence 26, Application: US/10062375
Patent No. US20020133000A1
GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.
de Lecea, Luis

Siggrins, George R.

Henriksen, Steven J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 NO. US20020133000A1th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/062.375

FILING DATE: 30-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/857.389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-10-062-375-26

Query Match 13.5%; Score 80.5; DB 12; Length 105;
Best Local Similarity 27.7%; Pred. No. 0.07;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLQALAAALPHRSVGSERWTFYPCPLR-----WRPKVKGPOL-KAK 81
DB 11 LLGATATAALPLEGGTGRDSEHMQEAAGIRKSSLLTFLAWFWFNTSQASAGPLIGEEA 70
QY 82 EDLRSVDNLP-----RERKAGCKNFYKGTSC 111
DB 71 REVARREGAPQOSARRDRMPCRNFFWKTSSC 104

RESULT 11

US-10-066-500-132

Sequence 132, Application: US/10066500

Patent No. US20020177165A1

GENERAL INFORMATION:

APPLICANT: Avi J. Ashkenazi

APPLICANT: Kevin P. Baker

APPLICANT: David A. Botstein

APPLICANT: Luc Desnoyers

APPLICANT: Dan L. Eaton

APPLICANT: Napoleone Ferrara

APPLICANT: Sherman Fong

APPLICANT: Wei-Qiang Gao

APPLICANT: Hanspeter Gerber

APPLICANT: Mary E. Gerritsen

APPLICANT: Audrey Goddard

APPLICANT: Paul J. Godowski

APPLICANT: Austin L. Gurney

APPLICANT: Ivar J. Kijavlin

APPLICANT: Jennie P. Mather

APPLICANT: Mary A. Napier

APPLICANT: James Pan

APPLICANT: Nicholas F. Paoni

APPLICANT: Margaret Ann Roy

APPLICANT: Timothy A. Stewart

APPLICANT: Daniel Tumas

APPLICANT: Colin K. Watanabe

APPLICANT: P. Mickey Williams

APPLICANT: William I. Wood

APPLICANT: Zemin Zang

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3130RIC7

CURRENT APPLICATION NUMBER: US/10/066,500

FILE FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 10/002,796

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/063329

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063733

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/066364

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066840

PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match

13.5% Score 80.5; DB 9; Length 155;

Best Local Similarity 27.78; Pred. No. 0.11;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLORALA-AALPHRSGVSRWRFPNCFLR-----WRPRKVKGPOL-KAK 81
DB 61 LLSGATAAALLEGPTGGDSEHMQAGIRKSSLLTFLAWFEWTSQASAGPLIGEEA 120
QY 82 EDLERSVDNLP-----RRKAGCKNFYWGFTSC 111
DB 121 REVAROEGAPPQOSARDRMPCRNFMTFSSC 154

RESULT 12

US-10-028-072-380
Sequence 380, Application US/10028072
Publication No. US20030004311A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang

TITLE OF INVENTION:

FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/049911

PRIOR FILING DATE: 1997-06-18

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059113

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059117

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059122

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059184

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059352

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 60/059836

PRIOR FILING DATE: 1997-09-24

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062287

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062814

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063045

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15

PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 13.5%; Score 80.5; DB 9; Length 155;
Best Local Similarity 27.7%; Pred. No. 0.11; Indels 19; Gaps 4;
Matches 26; Conservative 12; Mismatches 37;

QY 37 LLORALA-AALPHRSGVSEWRFTYPCNCLR-----WRPRKVGQOL-KAK 81
DB 61 LLGATATAALPLGGPTGRDSEHMOEAAAGIRKSSLLTFLAWFFWTSQASAGPLIGEEA 120
QY 82 EDLERSVDNLP-----RERKAGCKNFYWKGFYSC 111
DB 121 REVARROEGAPPOQSARRDRMPCRNFFWTFSSC 154
RESULT 13
US-10-121-049-380
Sequence 380, Application US/10121049
Publication No. US2003002239A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 380
LENGTH: 155
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-380
Query Match 13.5%; Score 80.5; DB 9; Length 155;
Best Local Similarity 27.7%; Pred. No. 0.11; Indels 19; Gaps 4;
Matches 26; Conservative 12; Mismatches 37;

QY 37 LLORALA-AALPHRSGVSEWRFTYPCNCLR-----WRPRKVGQOL-KAK 81
DB 61 LLGATATAALPLGGPTGRDSEHMOEAAAGIRKSSLLTFLAWFFWTSQASAGPLIGEEA 120
QY 82 EDLERSVDNLP-----RERKAGCKNFYWKGFYSC 111
DB 121 REVARROEGAPPOQSARRDRMPCRNFFWTFSSC 154

RESULT 14
US-10-123-904-380
Sequence 380, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven

APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 380
LENGTH: 155
TYPE: PRT
ORGANISM: Homo Sapien
S-10-123-904-380

Query Match 13.5%; Score 80.5; DB 9; Length 155;
Best Local Similarity 27.7%; Pred. NO. 0.11;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLORALA-AALPHRSGVSEWRTFFPNCPLR-----WRPKVKGQPL-KAK 81
DB 61 LLSGATATAALPLEGGPTGRDSEHMQEAAAGIRKSSLLTFLAWWFETISOASAGPLIGEAA 120
QY 82 EDLERSVDNLP-----RERKAGCKNFYWKGTSC 111
DB 121 REVARROEGAPPOQSARRDRMPCRNFFWTFSSC 154

RESULT 15

US-10-140-470-380
Sequence 380, Application US/10140470
Publication No. US20030022331A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 380
LENGTH: 155
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-380

Query Match 13.5%; Score 80.5; DB 9; Length 155;
Best Local Similarity 27.7%; Pred. NO. 0.11;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLORALA-AALPHRSGVSEWRTFFPNCPLR-----WRPKVKGQPL-KAK 81
DB 61 LLSGATATAALPLEGGPTGRDSEHMQEAAAGIRKSSLLTFLAWWFETISOASAGPLIGEAA 120

DB 61 LLSGATATAALPLEGGPTGRDSEHMQEAAAGIRKSSLLTFLAWWFETISOASAGPLIGEAA 120
QY 82 EDLERSVDNLP-----RERKAGCKNFYWKGTSC 111
DB 121 REVARROEGAPPOQSARRDRMPCRNFFWTFSSC 154

Search completed: March 21, 2003, 11:56:28
Job time : 23.4378 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:41:02 ; Search time 20.9614 Seconds
(without alignments)
155.808 Million cell updates/sec

Title: US-09-727-739B-15

Perfect score: 57
Sequence: 1 MRVSQTHCALGLGALAIC.....PPRRKAGCKNFYWGFTSC 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents:AA*

- 1: /cgn2_6ptcodata/2/1aa/5A.COMB.pep.*
- 2: /cgn2_6ptcodata/2/1aa/5B.COMB.pep.*
- 3: /cgn2_6ptcodata/2/1aa/6A.COMB.pep.*
- 4: /cgn2_6ptcodata/2/1aa/6B.COMB.pep.*
- 5: /cgn2_6ptcodata/2/1aa/PCRTUS.COMB.pep.*
- 6: /cgn2_6ptcodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160.5	26.9	110	3	US-08-648-322-3
2	87	14.6	112	3	US-08-648-322-2
3	87	14.6	112	4	US-09-001-472-3
4	83	13.9	85	3	US-08-648-322-6
5	80.5	13.5	105	4	US-09-001-472-2
6	79.5	13.3	109	3	US-08-648-322-5
7	78	13.1	29	3	US-08-648-322-7
8	78	13.1	29	3	US-08-648-322-11
9	78	13.1	94	3	US-08-648-322-10
10	77.5	13.0	943	1	US-08-455-970A-10
11	77	12.9	14	1	US-07-977-628A-1
12	77	12.9	14	1	US-08-253-272-5
13	77	12.9	14	1	US-08-416-007-4
14	77	12.9	14	1	US-08-676-263-11
15	77	12.9	14	1	US-08-286-748B-13
16	77	12.9	14	1	US-08-690-090A-1
17	77	12.9	14	2	US-08-488-159-1
18	77	12.9	14	2	US-08-465-764-1
19	77	12.9	14	2	US-08-475-751-4
20	77	12.9	14	2	US-08-283-980B-1
21	77	12.9	14	2	US-08-747-137-13
22	77	12.9	14	2	US-09-039-062-1
23	77	12.9	14	2	US-09-042-224-1
24	77	12.9	14	2	US-09-042-315A-1
25	77	12.9	14	3	US-08-931-095-1
26	77	12.9	14	3	US-09-100-414B-83
27	77	12.9	14	4	US-09-420-866-1

28	77	12.9	14	4	US-09-420-865-1	Sequence 1, Appl
29	77	12.9	14	4	US-09-303-323-83	Sequence 83, Appl
30	77	12.9	14	4	US-08-586-670A-1	Sequence 1, Appl
31	77	12.9	14	4	US-09-120-237-1	Sequence 1, Appl
32	77	12.9	14	4	US-09-144-931B-1	Sequence 1, Appl
33	77	12.9	14	4	US-09-331-405-1	Sequence 1, Appl
34	77	12.9	14	4	US-09-196-259-1	Sequence 1, Appl
35	77	12.9	14	6	5212156-1	Patent No. 5212156
36	77	12.9	14	6	5268278-4	Patent No. 5268278
37	77	12.9	15	6	5268278-1	Patent No. 5268278
38	77	12.9	31	3	US-09-100-414B-84	Sequence 84, Appl
39	77	12.9	31	3	US-09-100-414B-85	Sequence 85, Appl
40	77	12.9	31	3	US-09-100-414B-86	Sequence 86, Appl
41	77	12.9	31	3	US-09-100-414B-87	Sequence 87, Appl
42	77	12.9	31	3	US-09-100-414B-88	Sequence 88, Appl
43	77	12.9	31	3	US-09-100-414B-89	Sequence 89, Appl
44	77	12.9	31	4	US-09-303-323-84	Sequence 84, Appl
45	77	12.9	31	4	US-09-303-323-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-08-648-322-3
Sequence 3, Application US/08648322
Patent No. 6074872
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Gregor J.
TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-3

Query Match 26.9%; Score 160.5; DB 3; Length 110;
Best Local Similarity 38.3%; Pred No. 1e-12;
Matches 46; Conservative 17; Mismatches 30; Indels 27; Gaps 6;
QY 8 CALALGLGALAICS-OCGASOPDLASRLLORALAAALPHRSVGSVERWTFYFNCPL 66
DB 2 CALAALCIVLALGVGTGAPSDPRL-----RQFLKSLAAA-----TKGELAKYFLAE---L 50

QY 67 RWRPRKVGPKAKE-----DLERSVDNLP---PRRKAGCKNFYKGFSC 111
Db 51 LSEPNQENDALEFEDLPQAEQDEMRLEQSRANSANPAMAPRRKAGCKNFYKGFSC 110

RESULT 2

US-08-648-322-2

Sequence 2, Application US/08648322

Patent No. 6074872

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/648,322

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: 519.0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 554-2937

TELEFAX: (619) 554-6312

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 112 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-648-322-2

Query Match 14.6%; Score 87; DB 3; Length 112;

Best Local Similarity 25.2%; Pred. No. 0.0018;

Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSOGAASQPDLDASRRLLQALAAALPHRSVGER-----WRTF 59

4 CSTRGKRPSALSLLLLLLGGIAASALPLESGPTGQDSVQDATGGRRTGLLFLAWH-- 61

QY 60 YPNCPCLWRPRKVG-----POLKAKEDLERSVDNLPERRKAGCKNFYKGFSC 111

Db 62 -----EWASQDSSSTAEGTPELSKQ--ERPPLQPPHDKKPKCNFKWTFSSC 111

RESULT 3

US-09-001-472-3

Sequence 3, Application US/09001472

Patent No. 6232100

GENERAL INFORMATION:

APPLICANT: OLSEN, HENRIK S.

TITLE OF INVENTION: CORTISTATIN: POLYPEPTIDES

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: US
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001,472
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,980
FILING DATE: 31-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/037,386
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0430002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-001-472-3

Query Match 14.6%; Score 87; DB 4; Length 112;

Best Local Similarity 25.2%; Pred. No. 0.0018;

Matches 30; Conservative 9; Mismatches 42; Indels 38; Gaps 4;

QY 20 CSOGAASQPDLDASRRLLQALAAALPHRSVGER-----WRTF 59

Db 4 CSTRGKRPSALSLLLLLLGGIAASALPLESGPTGQDSVQDATGGRRTGLLFLAWH-- 61

QY 60 YPNCPCLWRPRKVG-----POLKAKEDLERSVDNLPERRKAGCKNFYKGFSC 111

Db 62 -----EWASQDSSSTAEGTPELSKQ--ERPPLQPPHDKKPKCNFKWTFSSC 111

RESULT 4

US-08-648-322-6

Sequence 6, Application US/08648322

Patent No. 6074872

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No. 6074872th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/648,322

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
US-08-648-322-6

Query Match 13.9%; Score 83; DB 3; Length 85;
Best Local Similarity 44.4%; Pred. No. 0.0041;
Matches 16; Conservative 5; Mismatches 13; Indels 2; Gaps 1;

QY 76 POLKAKEDLERSVDNLPPIRRKAGCKNYWKGFTSC 111
DB 51 PELSKRQ--ERPPLQPPHDKKPKCKNFFWKTSSC 84

RESULT 5

US-09-001-472-2
; Sequence 2, Application US/09001472
; Patent No. 6232100
; GENERAL INFORMATION:
; APPLICANT: OLSEN, HENRIK S.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: CORTISTATIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/001.472
FILING DATE: Herewith

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,980
FILING DATE: 31-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/037,386
FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0430002
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-001-472-2

Query Match 13.5%; Score 80.5; DB 4; Length 105;
Best Local Similarity 27.7%; Pred. No. 0.011;
Matches 26; Conservative 12; Mismatches 37; Indels 19; Gaps 4;

QY 37 LLORALA-AALPHRSGVSRWTFYPNCPCLR-----WPRKVKGPOL-KAK 81
DB 11 LLSGATATAALPLEGGTGDSEHQEAAGIRKSLTLFLAWFFWTSQASAGPLIGEA 70
QY 82 EDLERSVDNLP-----RERKAGCKNYWKGFTSC 111
DB 71 REVARROEGAPPOOSARRDRMPCRNFFWKTSSC 104

RESULT 6

US-08-648-322-5
; Sequence 5, Application US/08648322
; Patent No. 6074872
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecea, Luis
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/648,322
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: 519.0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 554-2937
TELEFAX: (619) 554-6312
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-648-322-5

Query Match 13.3%; Score 79.5; DB 3; Length 109;
Best Local Similarity 31.2%; Pred. No. 0.016;
Matches 30; Conservative 4; Mismatches 35; Indels 27; Gaps 4;

QY 37 LLORALAALPHRSG-----VSRWR-----TFYPCPLRW-----RPRKVKG 75
DB 19 LLGWVAASALPLESGFTGDSVQEATGSRGLLP-----LAWWHWASQASSSTPVGG 72
QY 76 POLKAKEDLERSVDNLPPIRRKAGCKNYWKGFTSC 111
DB 73 GTPGLSKSQRPPPPPPHLDKPKCKNFFWKTSSC 108

RESULT 7

US-08-648-322-7
; Sequence 7, Application US/08648322
; Patent No. 6074872
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecea, Luis
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 24

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
;; STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
;; CITY: La Jolla
;; STATE: California
;; COUNTRY: US
;; ZIP: 92037
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/648,322
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fitting, Thomas
;; REGISTRATION NUMBER: 34,163
;; REFERENCE/DOCKET NUMBER: 519.0
;; TELEPHONE: (619) 554-2937
;; TELEFAX: (619) 554-6312
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: C-terminal
;;
;; US-08-648-322-7

Query Match 13.1%; Score 78; DB 3; Length 29;
Best Local Similarity 51.9%; Pred. No. 0.004;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111
|| || | ||||| || ||
Db 2 ERPPQPPHLDKPKCKNFWKTFSSC 28

RESULT 8
US-08-648-322-11
; Sequence 11, Application US/08648322
; Patent No. 6074872
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecce, Luis
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/648,322
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: 519.0
; TELEPHONE: (619) 554-2937
; TELEFAX: (619) 554-6312
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: C-terminal
; US-08-648-322-10

Query Match 13.1%; Score 78; DB 3; Length 84;
Best Local Similarity 51.9%; Pred. No. 0.017;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111
|| || | ||||| || ||
Db 57 ERPPQPPHLDKPKCKNFWKTFSSC 83

RESULT 10
US-08-455-970A-10
; Sequence 10, Application US/08455970A

;; TELEFAX: (619) 554-6312
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: C-terminal
;;
;; US-08-648-322-11

Query Match 13.1%; Score 78; DB 3; Length 29;
Best Local Similarity 51.9%; Pred. No. 0.004;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111
|| || | ||||| || ||
Db 2 ERPPQPPHLDKPKCKNFWKTFSSC 28

RESULT 9
US-08-648-322-10
; Sequence 10, Application US/08648322
; Patent No. 6074872
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecce, Luis
; TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 NO. 6074872th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/648,322
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: 519.0
; TELEPHONE: (619) 554-2937
; TELEFAX: (619) 554-6312
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: C-terminal
; US-08-648-322-10

Query Match 13.1%; Score 78; DB 3; Length 84;
Best Local Similarity 51.9%; Pred. No. 0.017;
Matches 14; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 85 ERSVDNLPPEKAGCKNFYWKGTSC 111
|| || | ||||| || ||
Db 57 ERPPQPPHLDKPKCKNFWKTFSSC 83

RESULT 10
US-08-455-970A-10
; Sequence 10, Application US/08455970A

SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-255-272-5

Query Match 12.9%; Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
DB 1 AGCKNFYWKGTSC 14

RESULT 13

US-08-416-007-4
Sequence 4, Application US/08416007
Patent No. 5693679

GENERAL INFORMATION:

APPLICANT: Vincent, Jean-Pierre
APPLICANT: Gaudriault, Georges
APPLICANT: Beaudet, Alain
TITLE OF INVENTION: FLUORESCENT SOMATOSTATIN
NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416.007
FILING DATE: 04-APR-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06942/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-007-4

Query Match 12.9%; Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
DB 1 AGCKNFYWKGTSC 14

RESULT 14

US-08-676-263-11
Sequence 11, Application US/08676263

Patent No. 5705143
GENERAL INFORMATION:
APPLICANT: Bower, Gary R.
APPLICANT: Forster, Alan M.
APPLICANT: Riley, Anthony L. M.
APPLICANT: Storey, Anthony E.
TITLE OF INVENTION: BIOLOGICAL TARGETING AGENTS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676.263
FILING DATE: 07-NOV-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94300224.6
FILING DATE: 12-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Jeffrey S.
REGISTRATION NUMBER: 31,879
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: (312) 474-6600
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Cross-links
LOCATION: 3..14
US-08-676-263-11

Query Match 12.9%; Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 98 AGCKNFYWKGTSC 111
DB 1 AGCKNFYWKGTSC 14

RESULT 15

US-08-286-748B-13
Sequence 13, Application US/08286748B
Patent No. 5759542

GENERAL INFORMATION:

APPLICANT: Victor Gurevich
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,748B
FILING DATE: August 5, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: J. Peter Fasse
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04547/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-748B-13

Query Match 12.9% Score 77; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.002;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 98 ACCKNFYKGTSC 111
Db 1 ACCKNFYKGTSC 14

Search completed: March 21, 2003, 11:42:29
Job time: 21.9614 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```

: : March 21, 2003, 11:38:15 ; Search time 42.0773 Seconds
: : (without alignments)
: : 421.131 Million cell updates/sec

```

US-09-727-739B-17
450
Title: . . .
Perfect score:
Sequence: 1 MRVSIHCALALLGLALAIC.....RWRPRKVKGPQKAKEDLER 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

✓ Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45

```

Database : SPTREMBL_21:*

```

1:  sp_archea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_protein:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	136	30.2	120	13	Q90Y39	catostomus
2	112	24.9	115	13	Q90Y43	osteoglossu
3	108	24.0	114	13	Q90Y42	pantodon bu
4	93.5	20.8	114	13	Q90Y41	gnathomemus
5	88.5	19.7	116	13	Q90XEL	acipenser t
6	80.5	17.9	114	13	Q90Y40	chitala chi
7	78.5	17.4	603	16	Q88XA0	raistoncia s
8	76	16.9	251	16	Q88HG9	rhizobium l
9	68.5	15.2	435	16	Q9PAM4	xytelia fas
10	68	15.1	421	10	Q93W36	arabidopsis
11	68	15.1	578	10	Q8S8B0	arabidopsis
12	68	15.1	755	10	Q8S896	arabidopsis
13	68	15.1	758	10	Q82440	arabidopsis
14	68	15.1	772	10	Q9ZSP5	arabidopsis
15	67.5	15.0	197	16	Q8ZMR3	salmonella
16	67.5	15.0	568	16	Q8X1B0	raistoncia s

17	67	14.9	210	16	Q9HZ37	Q9HZ37 pseudomonas
18	67	14.9	210	1	Q9BSE9	Q9BSE9 homo sapien
19	67	14.9	338	16	Q9EPV4	Q9EPV4 streptomyce
20	67	14.9	698	10	Q89TE8	Q89TE8 oryza sativ
21	66.5	14.8	245	16	Q9BAM3	Q9BAM3 rhizobium l
22	66.5	14.8	252	16	Q9BAM3	Q9BAM3 neisseria m
23	66.5	14.8	523	10	Q9JRP3	Q9JRP3 oryza sativ
24	66.5	14.8	571	4	Q9JRC2	Q9JRC2 oryza sativ
25	66.5	14.8	753	4	Q9U776	Q9U776 homo sapien
26	66	14.7	789	10	Q9JRF2	Q9JRF2 homo sapien
27	65.5	14.6	201	2	Q54501	Q54501 streptomyce
28	65	14.4	144	16	Q9UHK3	Q9UHK3 homo sapien
29	65	14.4	397	16	Q9KZ43	Q9KZ43 streptomyce
30	64.5	14.3	201	2	Q87962	Q87962 streptomyce
31	63.5	14.1	678	10	Q9XF29	Q9XF29 oryza sativ
32	63.5	14.1	1060	15	Q97387	Q97387 chimpanzee
33	63	14.0	206	11	Q9EPQ4	Q9EPQ4 mus musculu
34	63	14.0	391	10	Q9ALG6	Q9ALG6 oryza sativ
35	63	14.0	393	11	Q9BEP3	Q9BEP3 mus musculu
36	63	14.0	431	17	Q9HN51	Q9HN51 halobacteri
37	63	14.0	486	10	Q9SBX3	Q9SBX3 zea mays (m
38	63	14.0	486	10	Q9FER7	Q9FER7 zea mays (m
39	63	14.0	723	2	Q97834	Q97834 pseudomonas
40	63	14.0	970	10	Q9MGD8	Q9MGD8 oryza sativ
41	62.5	13.9	145	3	Q9PA66	Q9PA66 letharia vu
42	62.5	13.9	295	16	Q8ZPH5	Q8ZPH5 salmonella
43	62.5	13.9	425	2	Q9AG14	Q9AG14 pseudomonas
44	62.5	13.9	531	16	Q9HVR9	Q9HVR9 pseudomonas
45	62	13.8	258	11	Q9RLI5	Q9RLI5 mus musculu

ALIGNMENTS

RESULT 1

Q90Y39		PRELIMINARY:	PRT:	120 AA..
ID	Q90Y39			
AC	Q90Y39			
DT	01-DEC-2001	(TREMBLrel. 19, Created)		
DD	01-DEC-2001	(TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)		
DE	Protoposomatostatin.			
OS	Catostomus commersoni (White sucker).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinodontiformes;			
OX	Catostomidae; Catostomus.			
NCBI_Taxid=7971;				
[1]	SEQUENCE FROM N.A.			
Al-Mahrouki A.A., Irwin D.M., Youson J.H.;				
"Molecular cloning and characterization of white sucker protposomatostatin";				
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.				
EMBL: AF293654; AAK97071.1; -				
Inter-Pro: IPR004250; Somatosatin.				
Pfam: PF03002; Somatosatin. 1.				
SEQUENCE 120 AA; 13783 MW; 00828032563E8805 CRC64;				

RESULT 2

Q90Y43


```

OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OX Osteoglossiformes; Notopteridae; Chitrala.
RN NCBI_TaxID=112163;
RP [1]
RQ SEQUENCE FROM N.A.
RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossiforms: molecular identification and comparative
RT analysis."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF292653; AAK97070.1;
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F5BE34F971 CRC64;
Query Match 17.9%; Score 80.5; DB 13; Length 114;
Best Local Similarity 50.0%; Pred. No. 0.13;
Matches 22; Conservative 6; Mismatches 13; Indels 3; Gaps 1;
QY 1 MRVSIOHCALALGALAIACSGAASQPDLDLASRRLLQALAA 44
Db 1 MLSTRIQCALLSLALPVSSTVAPSS---DLKLQLQRSIIA 41
RESULT 7
Q8XXA0 PRELIMINARY; PRT; 603 AA.
AC Q8XXA0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Probable DNA primase protein (EC 2.7.7.-)
GN DNAG OR RSC2216 OR RS01383.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RQ SEQUENCE FROM N.A.
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:457-502(2002)
DR EMBL; AL646068; CAD15923.1;
DR InterPro; IPR002936; DNAPrim_toprim.
DR Pfam; PF01751; Tprim_1.
DR Pfam; PF01807; Tprim_1.
DR Pfam; PF002688; Znf_CHC2; 1.
DR Pfam; PF002688; Znf_CHC2; 1.
DR SMART; SM00493; TOPRIM; 1.
DR Transferrase; Nucleotidyltransferase; Complete proteome.
SQ SEQUENCE 603 AA; 66305 MW; 5A3A29A1674E541C CRC64;
Query Match 17.4%; Score 78.5; DB 16; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.2;
Matches 22; Conservative 3; Mismatches 15; Indels 17; Gaps 1;
QY 9 ALALGLALAIACSGAASQPDLDLASRRLLQALAAALPH 48
Db 278 ALAQLGFANAVATLGTGVHVQKLLQVDFVIFSGDAGRAARRALEACPLH 334
RESULT 8
Q98HG9 PRELIMINARY; PRT; 251 AA.
AC Q98HG9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein ml12874.
GN ML12874.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RQ SEQUENCE FROM N.A.
RA Kato T., Sasamoto S.,
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohata M., Katsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Yabuta S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003000; BAB49897.1;
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 251 AA; 27987 MW; 45709C6A71EDD2AA CRC64;
Query Match 16.9%; Score 76; DB 16; Length 251;
Best Local Similarity 33.3%; Pred. No. 0.97;
Matches 29; Conservative 15; Mismatches 23; Indels 20; Gaps 5;
QY 3 VSQIHCALALGALAIACSGAASQPDLDLASRRLLQALAAALPHRSVSRWRTFYN 62
Db 180 VSALTREGLIALKIA---RVHYRGDLDAIAFL---AVAAAPKRR--SEAWR----- 226
QY 63 CCLRWLR-----PRKVGKPOLKAKEDL 84
Db 227 --CLRYRLKLAVERRLSAPKLEQSVL 251
RESULT 9
Q9PAM4 PRELIMINARY; PRT; 435 AA.
AC Q9PAM4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Xf2472.
GN Xf2472.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RQ SEQUENCE FROM N.A.
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.E., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.F.O., Netto L.E.S.,

```

RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.W., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 DR EMBL: AF004055; AF85270.1;
 DR InterPro: IPR002604; ATZ_TRZ.
 DR Pfam: PF01695; ATZ_TRZ.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 435 AA; 47299 MW; E527E57460CB447 CRC64;

Query Match 15.28; Score 68.5; DB 16; Length 435;
 Best Local Similarity 33.98; Pred. No. 12;
 Matches 20; Conservative 6; Mismatches 18; Indels 15; Gaps 1;

Qy 3 VSOIHC-----ALALLGLAICSGQAASQPDLDLASRLRLQALAAAL 46
 267 VSVVHCPSNKLASGFCFALHAGVNLAICTGDCASNLDMLSENRIAMLAARAV 325

RESULT 10
 ID Q93W36 PRELIMINARY; PRT; 421 AA.
 AC Q93W36;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE AT2G04160/T16923.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Miranda H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones,"
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [2]

SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones,"
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AY058235; AAL15409.1;
 DR EMBL: AY045647; AAK74005.1;
 DR InterPro: IPR003137; PA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF02225; PA: 1.
 DR Pfam: PF00082; Peptidase_S8: 1.
 DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN.1.
 SQ SEQUENCE 421 AA; 44863 MW; 5BB5AC7DAC436846 CRC64;

Query Match 15.18; Score 68; DB 10; Length 421;
 Best Local Similarity 28.18; Pred. No. 13;
 Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

Qy 14 GLAALICSGQAASQPDLDLASRLRLQALAA-----ALPHRSGVSEWRRTFYPCPLWRP 70
 178 GSVIAATYTGAVSPTNEQDFPRLLEFNAISGTSMSCPHISGIAGLLKTRYP-----WSP 232

Qy 71 RKVK 74
 Db 233 AAIR 236

RESULT 11
 ID Q8S8B0 PRELIMINARY; PRT; 578 AA.
 AC Q8S8B0;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Subtilisin-like serine protease AIR3 (Fragment).
 GN AT2G04160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007178; AAM15440.1;
 KW Protease.
 FT NON-TER
 SQ SEQUENCE 578 AA; 60733 MW; AFCAAPF484635E1C CRC64;

Query Match 15.18; Score 68; DB 10; Length 578;
 Best Local Similarity 28.18; Pred. No. 18;
 Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

Qy 14 GLAALICSGQAASQPDLDLASRLRLQALAA-----ALPHRSGVSEWRRTFYPCPLWRP 70
 335 GSVIAATYTGAVSPTNEQDFPRLLEFNAISGTSMSCPHISGIAGLLKTRYP-----WSP 389

Qy 71 RKVK 74
 Db 390 AAIR 393

RESULT 12
 ID Q8S896 PRELIMINARY; PRT; 755 AA.
 AC Q8S896;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Subtilisin-like serine protease AIR3 (Fragment).
 GN AT2G04160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,

```

RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007293; AAM15483.1;
KW Protease.
FT NON_TER 755 755
SQ SEQUENCE 755 AA; 80844 MW; 2E9BE365795936DE CRC64;

Query Match 15.1%; Score 68; DB 10; Length 758;
Best Local Similarity 28.1%; Pred. No. 24;
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

QY 14 GLALAIQSQAASQPDLDLASRLQALAA---ALPHRSGVSRWTFYPNCPCLRWRP 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 529 GVSVAAYTGAVSPTNEQDFPRLLFNALSGTSMSCPHISGIAGLLKTRYP-----WSP 583

QY 71 RKVK 74
   ||
DB 584 AAIR 587

RESULT 13
O82440 PRELIMINARY; PRT; 758 AA.
AC O82440:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Subtilisin-like protease (fragment).
GN AIR3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99178779; PubMed=10080694;
RA Neuteboom L.W., Ng J.M.Y., Kuyper M., Clijdesdale O.R.,
RA Hooykaas P.J.J., van der Zaai B.J.;
RT "Isolation and characterization of cDNA clones corresponding with
RT mRNAs that accumulate during auxin-induced lateral root formation.";
PL Plant Mol. Biol. 39:273-287(1999).
DR EMBL; AF055848; AAC62611.1;
DR HSP; Q45670; IDBI.
DR InterPro; IPR003137; PA.
DR Pfam; PF02225; PA; 1.
DR PRINTS; PR00082; Peptidase_S8; 2.
DR PROSITE; PS00138; SUBTILISIN.
DR PROSITE; PS00138; SUBTILISIN_SER; UNKNOWN_1.
KW Protease.
FT NON_TER 1
SQ SEQUENCE 758 AA; 81233 MW; E1683D2AD6C0516C CRC64;

Query Match 15.1%; Score 68; DB 10; Length 758;
Best Local Similarity 28.1%; Pred. No. 24;
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

QY 14 GLALAIQSQAASQPDLDLASRLQALAA---ALPHRSGVSRWTFYPNCPCLRWRP 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 GVSVAAYTGAVSPTNEQDFPRLLFNALSGTSMSCPHISGIAGLLKTRYP-----WSP 569

QY 71 RKVK 74
   ||
DB 570 AAIR 573

RESULT 14
O92SP5 PRELIMINARY; PRT; 772 AA.
ID O92SP5
AC O92SP5:
DT 01-MAY-1999 (TReMBLrel. 10, Created)

```

```

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Subtilisin-like protease.
GN AIR3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99246056; PubMed=10231025;
RA Neuteboom L.W., Veth-Tello L.M., Clijdesdale O.R., Hooykaas P.J.,
RA van der Zaai B.J.;
RT "A novel subtilisin-like protease gene from Arabidopsis thaliana is
RT expressed at sites of lateral root emergence.";
RL DNA Res. 6:13-19(1999).
DR EMBL; AF098632; AAD12260.1;
DR HSP; Q45670; IDBI.
DR InterPro; IPR003137; PA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02225; PA; 1.
DR PRINTS; PR00082; Peptidase_S8; 2.
DR PROSITE; PR00723; SUBTILISIN.
DR PROSITE; PS00138; SUBTILISIN_SER; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 772 AA; 82873 MW; 75DCCA0DED63F47E CRC64;

Query Match 15.1%; Score 68; DB 10; Length 772;
Best Local Similarity 28.1%; Pred. No. 24;
Matches 18; Conservative 13; Mismatches 25; Indels 8; Gaps 2;

QY 14 GLALAIQSQAASQPDLDLASRLQALAA---ALPHRSGVSRWTFYPNCPCLRWRP 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 529 GVSVAAYTGAVSPTNEQDFPRLLFNALSGTSMSCPHISGIAGLLKTRYP-----WSP 583

QY 71 RKVK 74
   ||
DB 584 AAIR 587

RESULT 15
O82MR3 PRELIMINARY; PRT; 197 AA.
ID O82MR3
AC O82MR3:
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative periplasmic protein.
GN STM2741.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008824; AAL21627.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 197 AA; 22326 MW; 7CA059444F962F66 CRC64;

Query Match 15.0%; Score 67.5; DB 16; Length 197;
Best Local Similarity 33.3%; Pred. No. 7;

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:35:23 ; Search time 10.3348 Seconds
(without alignments)
345.142 Million cell updates/sec

Title: US-09-727-739B-17

Perfect score: 450

Sequence: 1 MRVSOIHCALLGLALALIC.....RWRPRKVGPKLAKEDLER 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	347.5	77.2	115	1 SMS2_ONCMY	Q91194 oncorhynch
2	119	26.4	120	1 SMS2_CARAU	Q9Y9H4 carassius a
3	81.3	20.3	114	1 SMSA_CHICK	Q9Y9H5 carassius a
4	83.3	18.6	116	1 SMSI_ICTUPU	P33094 gallus gall
5	80.3	17.9	114	1 SMSI_HUMAN	P01171 ictalurus p
6	79.5	17.7	116	1 SMSI_HUMAN	P01166 homo sapien
7	76.5	17.0	115	1 SMSI_RANRI	P87384 rana ridibu
8	75.5	16.8	115	1 SMSI_PROAN	Q9W710 protopterus
9	69.5	15.4	116	1 SMS_SHEEP	O46688 ovis aries
10	69.5	15.4	125	1 SMS2_LOPAM	P01170 lophius ame
11	68.5	15.2	116	1 SMS_BOVIN	P26917 bos taurus
12	67.5	15.0	314	1 Y4RB_RHISN	P55635 rhizobium s
13	67	14.9	1204	1 NOS3_BOVIN	P29473 bos taurus
14	66	14.7	260	1 PT22_SACBA	O13374 saccharomyc
15	65.5	14.6	116	1 SMS_MOUSE	P01167 mus musculu
16	64.5	14.3	1305	1 GAK_RAT	P97874 rattus norv
17	63.5	14.1	116	1 SMS_CANFA	P49670 canis fami
18	63	14.0	1204	1 NOS3_PIG	Q28969 sus scrofa
19	62	13.8	308	1 ARCC_SYNY3	Q28969 sus scrofa
20	61.5	13.7	337	1 YAJ2_PSEAE	P74733 synechocyst
21	60.5	13.4	321	1 YNEE_ECO57	Q9HVA4 pseudomonas
22	60.5	13.4	469	1 CYAA_TRYEQ	O8XAZ3 escherichia
23	60.5	13.4	1235	1 CYAA_TRYEQ	P26338 trypanosoma
24	60.5	13.4	1310	1 CYAA_TRYEQ	Q26721 trypanosoma
25	59.5	13.2	321	1 CTM4_MOUSE	Q99A47 mus musculu
26	59.5	13.2	350	1 YNEE_ECOLI	P76146 escherichia
27	59.5	13.2	550	1 YLOO_MCTU	Q10709 mycobacteri
28	59	13.2	2476	1 ZAN_PIG	Q28983 sus scrofa
29	58.5	13.0	325	1 EGIP_ANTCR	P15217 anthocidari
30	58.5	13.0	396	1 DHH_HUMAN	O43323 homo sapien
31	58.5	13.0	1115	1 TBC2_CHLRE	O8VXP3 chlamydomon
32	58	12.9	1288	1 CTA3_HUMAN	O9BZ76 homo sapien
33	58	12.9	582	1 PRIM_PASMO	Q9C119 pasteurella
			720	1 PRDH_HUMAN	Q9H9d4 homo sapien

34 58 12.9 941 1 DNAB_RHOMR
35 58 12.9 1503 1 TRL2_HUMAN
36 57.5 12.8 894 1 POL2_BAMMU
37 57.5 12.8 988 1 E4L2_MOUSE
38 57.5 12.8 1077 1 E2K3_CABEL
39 57 12.7 397 1 UL50_HCMVA
40 57 12.7 479 1 ORC5_YEAST
41 57 12.7 834 1 MSH5_HUMAN
42 57 12.7 1042 1 SPAL_HUMAN
43 57 12.7 3674 1 SPCL_HUMAN
44 56.5 12.6 195 1 NIFQ_AZOVI
45 56.5 12.6 347 1 DLX4_HUMAN

ALIGNMENTS

RESULT 1
SMS2_ONCMY STANDARD: PRT: 115 AA.
AC Q91194;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
DE [Tyr7,Gly10]somatostatin-14].
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95354921; PubMed=7628684;
RX Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;
RT "Isolation and characterization of a cDNA encoding for
RT pro-somatostatin containing [Tyr7, Gly10]-somatostatin-14 from the
RT endocrine pancreas of rainbow trout, Oncorhynchus mykiss";
RL Gen. Comp. Endocrinol. 98:253-261(1995).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32471; AAC59695.1;
CC InterPro; IPR004250; Somatostatin;
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 87
FT PEPTIDE 88 115 [TYR21, GLY24]SOMATOSTATIN-28 (POTENTIAL).
FT PEPTIDE 102 115 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFD 104 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12963 MW; 520595025FCA6D91 CRC64;
Query Match 77.2%; Score 347.5; DB 1; Length 115;
Best Local Similarity 85.7%; Pred. No. 2.2e-32;
Matches 72; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 1 MRVSOIHCALLGLALALICSGAASQPDLDLSRRLLQRAAALPHRSGVSRWRTPY 60
DB 1 MRVSOIHCALLGLALALICSGAASQPDLDLSRRLLQRAAALPHRSGVSRWRTPY 60

QY 61 PNCPLRWRPRKVGPKLAKEDL 84
DB 61 PNCPLRWRPRKVGPKLAKEDL 84


```

RESULT 2
ID SMS2_CARAU STANDARD; PRT; 120 AA.
AC Q9YGH4: Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Somatostatin II precursor (Contains: [Tyr11,Gly24]somatostatin-28;
DE [Tyr7,Gly10]somatostatin-14);
OS Carassius auratus (Goldfish);
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
CC Cyprinidae; Carassius.
OX NCBI_TaxID=7937;
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I
RT and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U60262; AAD09626.1;
DR InterPro; IPR004250; Somatostatin.
KW Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL 1 23
FT PROPEP 24 92 POTENTIAL.
FT PEPTIDE 93 120 [TYR21, GLY24] SOMATOSTATIN-28.
FT PEPTIDE 107 120 [TYR7, GLY10] SOMATOSTATIN-14.
FT DISULFID 109 120 BY SIMILARITY.
FT CONFLICT 51 51 Q -> RW (IN REF. 2).
SQ SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;
Query Match. 26.4%; Score 119; DB 1; Length 120;
Best Local Similarity 57.8%; Pred. No. 1.3e-06;
Matches 26; Conservative 6; Mismatches 11; Indels 2; Gaps 1;
QY 1 MRVSQTHCALALGLALALCSQGAASQ--PDLDLASRRLLQRLA 43
DB 1 MRLCELHCYLALLGLSLVLCGRCAANSQLEPDLDLFRHRLQRLA 45
RESULT 3
ID SMSA_CARAU STANDARD; PRT; 114 AA.
AC Q9YGH5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Somatostatin IA precursor (Contains: Somatostatin-26; Somatostatin-
DE 14);
OS Carassius auratus (Goldfish).

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Carassius.
NCBI_TaxID=7937;
SEQUENCE FROM N.A.
TISSUE=Brain;
Lin X.-W., Peter R.E.;
Cloning and characterization of cDNAs encoding preprosomatostatin-I
and -II from goldfish brain.";
Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
-1- SUBCELLULAR LOCATION: Secreted.
-1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; U40734; AAD09359.1;
InterPro; IPR004250; Somatostatin.
Pfam; PF03002; Somatostatin; 1.
Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
SIGNAL 1 24 POTENTIAL.
PROPEP 25 88
PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).
PEPTIDE 101 114 SOMATOSTATIN-14.
DISULFID 103 114 BY SIMILARITY.
SEQUENCE 114 AA; 12574 MW; B5920015E2D272A4 CRC64;
Query Match. 20.3%; Score 91.5; DB 1; Length 114;
Best Local Similarity 54.8%; Pred. No. 0.0016;
Matches 23; Conservative 5; Mismatches 11; Indels 3; Gaps 1;
QY 1 MRVSQTHCALALGLALALCSQGAASQPDLDLASRRLLQRLA 42
DB 1 MLSTRIQCALALLSLALALCVSVSA--PTDAKLRLQLRSL 39
RESULT 4
ID SMS_CHICK STANDARD; PRT; 116 AA.
AC P33094;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor (Contains: Somatostatin-28; Somatostatin-14).
GN SST.
OS Gallus gallus (Chicken).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
CC Gallus.
OX NCBI_TaxID=9031;
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RA Nata K., Kobayashi T., Karahashi K., Kato S., Yamamoto H.,
RA Yonekura H., Okamoto H.;
RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


CC tumors and vasoactive intestinal peptide tumors.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J00306; AAA60566.1;
CC DR EMBL: M19318; AAA35908.1;
CC DR PIR: A43614; RHUS;
CC DR PIR: A28968; A28968.
CC DR Genew: HGNC:11329; SST.
CC DR MIM: 162450;
CC DR InterPro: IPR004250; Somatostatin.
CC DR Pfam: PF03002; Somatostatin; 1.
CC KW Cleavage on pair of basic residues; Hormone; Signal; Pharmaceutical:
FT SIGNAL 1 24
FT PROPEP 25 88
FT PEPTIDE 89 116 SOMATOSTATIN-28;
FT PEPTIDE 103 116 SOMATOSTATIN-14;
FT DISULFID 105 116
SQ SEQUENCE 116 AA; 12735 MW; AB49BB89DC9DD8DA.CRC64;

Query Match 17.7%; Score 79.5; DB 1; Length 116;
Best Local Similarity 34.9%; Pred. No. 0.037;
Matches 29; Conservative 13; Mismatches 28; Indels 13; Gaps 5;

QY 5 QHCAALGLALAI-CSGAAASOPDLGLASRLQLALAAALPHRSVGSFWRFTVPNC 63
DB 5 LQCALAALSIVLALGCVTPGSDPRL----RQFLQKSLAAA----AGKELAKYFLAE- 55
QY 64 PCLRRPRKVKGPOLKAKEDLER 86
DB 56 --LLSEPNQTENDALE--PEDLSQ 75

RESULT 7
ID SMSL_RANRI STANDARD; PRT; 115 AA.
AC P87384; Q9PS18;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-14 (S-1)
(SS1)].
OS Rana ridibunda (laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranolia; Rana.
NCBI_TaxID=8406;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Tostivint H., Lihrmann I., Bucharies C., Vieau D., Coulouarn Y.,
RA Fournier A., Conlon J.M., Vaudry H.;
RT "Occurrence of two somatostatin variants in the frog brain:
RT characterization of the cDNAs; distribution of the mRNAs, and
RT receptor-binding affinities of the peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
RN [2]
RP SEQUENCE OF 102-115.
RC TISSUE=Brain;
RX MEDLINE=93038702; PubMed=1358069;
RA Vaudry H., Charrel N., Conlon J.M.;
RT "Isolation of [Pro2, Met13]somatostatin-14 and somatostatin-14 from the
RT frog brain reveals the existence of a somatostatin gene family in a
RT tetrapod.";
RL Biochem. Biophys. Res. Commun. 188:477-482(1992).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U68136; AAC60093.1;
CC DR InterPro: IPR004250; Somatostatin.
CC DR Pfam: PF03002; Somatostatin; 1.
CC KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT PROPEP 25 99 BY SIMILARITY.
FT PEPTIDE 102 115 SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12691 MW; 349756FEB4BE213.CRC64;

Query Match 17.0%; Score 76.5; DB 1; Length 115;
Best Local Similarity 47.8%; Pred. No. 0.079;
Matches 22; Conservative 8; Mismatches 11; Indels 5; Gaps 2;

QY 1 MRVSIQHCAALGLALAI-CISGAA-SQPDLDLASRLRLQLALAAA 45
DB 1 MQSCRVOCALTLSSLALAINSAIAPTDPRL----RQFLQKSLASA 42

RESULT 8
ID SMSL_PROAN STANDARD; PRT; 115 AA.
AC Q9WTF0.2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-27;
DE Somatostatin-14].
OS Protopterus annectens (African lungfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Lepidosireniformes; Protopteriidae; Protopterus.
NCBI_TaxID=7888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99326690; PubMed=10398054;
RA Trabucchi M., Tostivint H., Lihrmann I., Jegou S., Vallarino M.,
RA Vaudry H.;
RT "Molecular cloning of the cDNAs and distribution of the mRNAs encoding
RT two somatostatin precursors in the African lungfish Protopterus
RT annectens.";
RL J. Comp. Neurol. 410:643-652(1999).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF126243; AAD39138.1;
CC DR InterPro: IPR004250; Somatostatin.
CC DR Pfam: PF03002; Somatostatin; 1.
CC KW Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 88 POTENTIAL.
FT PEPTIDE 89 115 SOMATOSTATIN-27 (POTENTIAL).
FT PEPTIDE 102 115 SOMATOSTATIN-14.

```

FT  DISULFID 104 115 BY SIMILARITY
SQ  SEQUENCE 115 AA; 12600 MW; B0CFE1E603FEAF09 CRC64;

Query Match 16.8%; Score 75.5; DB 1; Length 115;
Best Local Similarity 48.8%; Pred. No. 0.1; Indels 3; Gaps 1;
Matches 20; Conservative 4; Mismatches 14;

QY 5 QIHCALALLGLALALICSGAASQPDLDLASRRLLQALAA 45
DB 5 RFOCALVLLSLAVFSKVSAAAPS---DLRLQLLQSLAA 42

RESULT 9
SMS_SHEEP
ID SMS_SHEEP STANDARD; PRT; 116 AA.
AC O46688;
DT 15-DEC-1998 (Rel. 37, Created)
JT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
GN SST.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ile de france;
RX MEDLINE=99094691; PubMed=9880082;
RA Bruneau G., Tillet Y.;
RT "Localization of the preprosomatostatin-mRNA by in situ hybridization
in the ewe hypothalamus.";
RL Peptides 19:1749-1758(1998).
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF031488; AAC04697.1;
CC EMBL; Y15267; CAAY5556.1;
CC InterPro; IPR004250; Somatostatin; 1.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Signal.
CC SIGNAL 1 24 BY SIMILARITY.
CC PROPEP 25 88 BY SIMILARITY.
CC PEPTIDE 89 116 SOMATOSTATIN-28.
CC PEPTIDE 103 116 SOMATOSTATIN-14.
CC DISULFID 105 116 BY SIMILARITY.
SQ SEQUENCE 116 AA; 12689 MW; C18F17E31A3718DE CRC64;

Query Match 15.4%; Score 69.5; DB 1; Length 116;
Best Local Similarity 33.7%; Pred. No. 0.49; Indels 13; Gaps 5;
Matches 28; Conservative 13; Mismatches 29;

QY 5 QIHCALALLGLALALICSGAASQPDLDLASRRLLQALAAALPHRSGVSEWRTPYNC 63
DB 5 RQCALAALSIVLALGVGTGAPSDPRL---RQFLQKSLAAA-----AGQELAKYFLAE- 55

QY 64 PCLRWPRKVGKPOLKAKEDLER 86
DB 56 ---LLSEPNQENDALE-PEDLSQ 75.

RESULT 10
SMS2_LOPAM
ID SMS2_LOPAM STANDARD; PRT; 125 AA.
AC P01170; G91066;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr7,Gly10]somatostatin-14].
OS Lophius americanus (American goosefish) (Anglerfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81052423; PubMed=6107860;
RA Hobart P.M., Crawford R., Shen L., Pictet R., Rutter W.J.;
RT "Cloning and sequence analysis of cDNAs encoding two distinct
somatostatin precursors found in the endocrine pancreas of
anglerfish.";
RL Nature 288:137-141(1990).
RN [2]
RP PARTIAL SEQUENCE, AND HYDROXYLATION.
RX MEDLINE=87308304; PubMed=2887572;
RA Andrews P.C., Nichols R., Dixon J.E.;
RT "Post-translational processing of preprosomatostatin-II examined
using fast atom bombardment mass spectrometry.";
RL J. Biol. Chem. 262:12692-12699(1987).
CC -!- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: SOMATOSTATIN II MAY HAVE A DIFFERENT DEGREE OF
ACTIVITY OR A DIFFERENT TYPE OF TARGET CELL FROM SOMATOSTATIN I.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; V00641; CAA23987.1;
CC PIR; A01434; RIAFS2.
CC PIR; A27376; A27376.
CC InterPro; IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Signal; Hydroxylation;
KW Multigene family.
KW SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 109
FT PEPTIDE 112 125 [TYR7, GLY10]SOMATOSTATIN-14.
FT DISULFID 114 125
FT MOD_RES 120 120 HYDROXYLATION
FT CONFLICT 77 78 DV -> TG (IN REF. 1).
FT CONFLICT 90 90 G -> E (IN REF. 1).
SQ SEQUENCE 125 AA; 14052 MW; 5E14605D7B9A46FE CRC64;

Query Match 15.4%; Score 69.5; DB 1; Length 125;
Best Local Similarity 36.1%; Pred. No. 0.53; Indels 15; Gaps 3;
Matches 22; Conservative 7; Mismatches 17;

QY 6 IHCALALLGLALALICSGAASQ-----PDLDASRR--LLQALAAALPHRSGVSEWR 55
DB 4 IRCFALLALLALVLCGFSVSSQDLREQSDNQDLDELRLQHWLLERARSAGL-----LSQE 58

QY 56 W 56
DB 59 W 59

RESULT 11
SMS_BOVIN
ID SMS_BOVIN STANDARD; PRT; 116 AA.
AC P26917;

```

DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
 GN SST
 OS Bos taurus (Bovine)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88288237; PubMed=2899837;
 RA Su C.J., White J.W., Li W.H., Luo C.C., Frazier M.L., Saunders G.F.,
 RA Chan L.;
 RT "Structure and evolution of somatostatin genes."
 RL Mol. Endocrinol. 2:209-216(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Holstein;
 RX MEDLINE=99198780; PubMed=10100681;
 RA Furu L.M., Kazner G.W., Strausbaugh L., Zinn S.A.;
 RT "Cloning and characterization of the bovine somatostatin gene."
 RL J. Anim. Sci. 77:492-493(1999).
 CC - FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M31217; AAA30744.1;
 DR EMBL; U97077; AAB58056.1;
 DR PIR; A40929;
 DR InterPro; IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Signal.
 FT SIGNAL 1 24 BY SIMILARITY.
 FT PROPEP 25 88 BY SIMILARITY.
 FT PEPTIDE 89 116 SOMATOSTATIN-28.
 FT PEPTIDE 103 116 SOMATOSTATIN-14.
 FT DISULFID 105 116 BY SIMILARITY.
 SQ SEQUENCE 116 AA; 12688 MW; C18F17E64A371D8E CRC64;
 Query Match 15.2%; Score 68.5; DB 1; Length 116;
 Best Local Similarity 45.2%; Pred. No. 0.64;
 Matches 19; Conservative 7; Mismatches 11; Indels 5; Gaps 2;
 YJ 5 QIHCALALLGLAICSGQASQPDLDLSRRLLQRLAALAA 45
 Db 5 RLQCALAALSIVLAIGVGTGAPSDPRL-----RQFLQKSLAA 42
 RESULT 12
 YARB_RHISN
 ID YARB_RHISN STANDARD; PRT; 314 AA.
 AC P55635;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative integrase/recombinase YARB.
 GN YARB.
 OS Rhizobium sp. (strain NGR234).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes."
 RL Nature 387:394-401(1997).
 CC - SIMILARITY: BELONGS TO THE "PHAGE" INTEGRASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AE000093; AAB92468.1;
 DR InterPro; IPR002104; Phage_integrase.
 DR Pfam; PF00589; Phage_integrase; 1.
 KW Hypothetical protein; DNA recombination; DNA integration; Plasmid;
 KW Transposable element.
 SQ SEQUENCE 314 AA; 36310 MW; 9D33FBESAAE95068 CRC64;
 Query Match 15.0%; Score 67.5; DB 1; Length 314;
 Best Local Similarity 33.3%; Pred. No. 2.3;
 Matches 25; Conservative 11; Mismatches 26; Indels 13; Gaps 3;
 QY 15 LALAICSGQASQPD---LDLSRRLLQRLAALAAALPHSGVSEWRWTFYPCNCLRWPR 71
 Db 48 LALAAQPPSPQAQPSHWATPLGVVRYRYRAADP-----RTQIPPSGLLPFRK 98
 QY 72 KVGQPOLKXKEDLER 86
 Db 99 RAR-PYLSKEDIQR 112
 RESULT 13
 NOS3_BOVIN
 ID NOS3_BOVIN STANDARD; PRT; 1204 AA.
 AC P29473;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nitric-oxide synthase, endothelial (EC 1.14.13.39) (EC-NOS) (NOS, type
 DE III) (NOSIII) (Endothelial NOS) (eNOS) (Constitutive NOS) (cNOS).
 GN NOS3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92335295; PubMed=1378626;
 RA Lamas S., Marsden P.A., Li G.K., Tempst P., Michel T.;
 RT "Endothelial nitric oxide synthase: molecular cloning and
 RT characterization of a distinct constitutive enzyme isoform."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6348-6352(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93055452; PubMed=1385480;
 RA Nishida K., Harrison D.G., Navas J.P., Fisher A.A., Dockery S.P.,
 RA Umatsu M., Nerem R.M., Alexander R.W., Murphy T.J.;
 RT "Molecular cloning and characterization of the constitutive bovine
 RT aortic endothelial cell nitric oxide synthase."
 RL J. Clin. Invest. 90:2092-2096(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=92348367; PubMed=1379225;
 RA Sessa W.C., Harrison J.K., Barber C.M., Zeng D., Durieux M.E.,
 RA D'Angelo D., Iyach K.R., Peach M.J.;
 RT "Molecular cloning and expression of a cDNA encoding endothelial cell


```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; K02248; AAA42161.1; -
DR EMBL; V01271; CAA24579.1; -
DR EMBL; J00787; AAA42164.1; -
DR EMBL; M25890; AAA42167.1; -
DR EMBL; J00788; AAA42162.1; -
DR EMBL; X51468; CAA35831.1; -
DR PIR; A20983; RIRTS1.
DR PIR; S08416; S08416.
DR MGD; MGI:98326; Smst.
DR InterPro; IPR004250; Somatostatin.
DR Pfam; PF03002; Somatostatin; 1.
KW Cleavage on pair of basic residues; Hormone; Signal.
FT SIGNAL 25 24 ANTRIN.
FT PEPTIDE 25 34
FT PROPEP 35 88
FT PEPTIDE 89 116 SOMATOSTATIN-28.
FT PEPTIDE 103 116 SOMATOSTATIN-14.
FT DISULFID 105 116
FT CONFLICT 43 43 T -> Y (IN REF. 5).
FT CONFLICT 79 79 Q -> H (IN REF. 6).
FT SEQUENCE 116 AA; 12745 MW; D48B5454C4490375 CRC64;
SQ
Query Match 14.69; Score 65.5; DB 1; Length 116;
Best Local Similarity 45.28; Pred. No. 1.4;
Matches 19; Conservative 7; Mismatches 11; Indels 5; Gaps 2;
QY 5 QIHCALALLGLALATCS-QGAASQPDLDLASRRLLQRLALAAA 45
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 RLQCALAALCIVLALGGVTGAPSPRL-----RFLQKSLAA 42

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:38:56 ; Search time 18.4549 Seconds
(without alignments)
447,986 Million cell updates/sec

Title: US-09-727-739b-17

Perfect score: 450

Sequence: 1 MRYQIHCALALGLALALIC.....RPRKVKGPQLKAKEDLR 86

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR-73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	347.5	77.2	115	2 I51064	somatostatin II pr
2	83.5	18.6	116	1 S20630	somatostatin precu
3	80.5	17.9	114	1 RIIDS1	somatostatin-14 pr
4	80.5	17.9	114	2 I50798	preprosomatostatin
5	79.5	17.7	116	1 RIHUS1	somatostatin I pre
6	79.5	17.7	116	1 A28968	somatostatin I pre
7	76.5	17.0	115	2 JC6166	somatostatin-14 pr
8	69.5	15.4	125	1 RIAPS2	somatostatin II pr
9	68.5	15.2	116	1 RIPOS1	somatostatin precu
10	68.5	15.2	435	2 A82554	conserved hypothet
11	68	15.1	758	2 T51335	subtilisin-like pr
12	67	14.9	120	2 F83244	hypothetical prote
13	67	14.9	1205	1 A38443	nitric-oxide synth
14	66.5	14.8	259	2 H81009	exodeoxyribonuclea
15	66.5	14.6	571	2 T43456	hypothetical prote
16	65.5	14.6	116	1 RIHUS1	somatostatin precu
17	65.5	14.6	116	1 RIHUS1	somatostatin precu
18	65.5	14.6	201	2 S55000	alpha-chitin bindi
19	64.5	14.3	1305	2 T31096	cyclin G-associate
20	63	14.0	431	2 F84375	N-ethylameline ch
21	62.5	13.9	531	2 B83082	probable binding p
22	62	13.8	308	1 S76941	carbamate kinase (
23	62	13.8	472	2 T36529	probable two-comp
24	62	13.8	501	2 AF0346	probable GntR-fam
25	61.5	13.7	256	2 AG0676	probable ATP/GTP-b
26	61.5	13.7	327	2 T34666	probable regulator
27	61.5	13.7	337	2 C83059	conserved hypothet
28	61.5	13.7	398	2 C81141	hypothetical prote
29	61.5	13.7	706	2 T01351	subtilisin-like pr

30	61	13.6	238	2 H83559	hypothetical prote
31	61	13.6	375	2 G81878	probable lipoprote
32	61	13.6	519	2 D70695	hypothetical prote
33	60.5	13.4	273	2 E75476	hypothetical prote
34	60.5	13.4	321	2 G90894	hypothetical prote
35	60.5	13.4	321	2 B85723	hypothetical prote
36	60.5	13.4	469	2 S16359	adenylate cyclase
37	60.5	13.4	474	2 C75625	hypothetical prote
38	60.5	13.4	1235	2 D32433	VSG expression sit
39	60	13.3	321	2 F75287	probable DNA polym
40	60	13.3	433	2 H86597	flagellum-specific
41	60	13.3	433	2 B72025	flagellum-specific
42	60	13.3	433	2 A81513	virulence APrase,
43	60	13.3	503	2 A83027	hypothetical prote
44	59.5	13.2	256	2 G86682	hypothetical prote
45	59.5	13.2	304	2 A83309	hypothetical prote

ALIGNMENTS

RESULT 1

I51064

somatostatin II precursor - rainbow trout

C:Species: Oncorhynchus mykiss (rainbow trout)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C:Accession: I51064

R:Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.

Gen. Comp. Endocrinol. 98, 253-261, 1995

A:Title: Isolation and characterization of a cDNA encoding for preprosomatostatin c

A:Reference number: I51064; MUID:95354921; PMID:7628684

A:Accession: I51064

A:Status: Preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-115 <MOO>

A:Cross-references: EMBL:U32471; NID:975344; PIDN:AAC59695.1; PID:975345

C:Superfamily: somatostatin

Query Match 77.2%; Score 347.5; DB 2; Length 115;

Best Local Similarity 85.7%; Pred. No. 9.6e-32;

Matches 72; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 1 MRYQIHCALALGLALALICSGAASQPDLDLRSRLRLQRLAALPHRSGVSRWRTFY 60

DB 1 MKVRIHCALALGLALALICSGAASQPDLDLRSRLRLQRLAALPHRSGVSRWRTFY 60

QY 61 PNCPLRWPRKVKGPQLKAKEDL 84

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

QY 61 PNCPLRWPRKVKGPQLKAKEDL 81

DB 61 PNCPLRWPRKVKGPQLKAKEDL 81

Db 5 RLOCALALLSIALVGTWVSAPDRL-----ROFLQKSLAAA-----AKQELAKYFLAE- 55
 QY 64 PCURWPRVKGPQKAKEDLER 86
 Db 56 --LLSEPSQTENEALLES-EDLSR 75

RESULT 3

RIHDSI
 somatostatin-14 precursor - channel catfish
 N:Alternate names: somatostatin I
 C:Species: Ictalurus punctatus (channel catfish)
 C:Date: 30-Jun-1980 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
 C:Accession: S00292; A93897; A92334; A01435
 R:Minth, C.D.; Taylor, W.L.; Magazin, M.; Tavianini, M.A.; Collier, K.; Weith, H.L.; Dixon, J.E. Biol. Chem. 257, 10372-10377, 1982
 A:Title: The structure of cloned DNA complementary to catfish pancreatic somatostatin-14
 A:Reference number: S00292; MUID:82265638; PMID:6179939
 A:Accession: S00292
 A:Molecule type: mRNA
 A:Residues: 1-114 <MIN>
 A:Cross-references: EMBL:V00607; NID:964017; PIDN:CAA23877.1; PID:964018
 Taylor, W.L.; Collier, K.J.; Deschenes, R.J.; Weith, H.L.; Dixon, J.E.
 Proc. Natl. Acad. Sci. U.S.A. 78, 6694-6698, 1981
 A:Title: Sequence analysis of a cDNA coding for a pancreatic precursor to somatostatin.
 A:Reference number: A93897; MUID:82082515; PMID:6171821
 A:Accession: A93897
 A:Molecule type: mRNA
 A:Residues: 82-108 <TAY>
 A:Cross-references: GB:J00944
 R:Andrews, P.C.; Dixon, J.E.
 J. Biol. Chem. 256, 8267-8270, 1981
 A:Title: Isolation and structure of a peptide hormone predicted from a mRNA sequence. A
 A:Reference number: A92334; A92334; MUID:81264223; PMID:6114953
 A:Accession: A92334
 A:Molecule type: protein
 A:Residues: 101-114 <AND>
 A:Superfamily: somatostatin
 C:Keywords: neuropeptide
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-100/Domain: propeptide #status predicted <PRO>
 F:101-114/Product: somatostatin-14 #status experimental <MAT>
 F:103-114/Disulfide bonds: #status experimental

Query Match 17.9%; Score 80.5; DB 1; Length 114;
 Best Local Similarity 46.7%; Pred. No. 0.078;
 Matches 21; Conservative 8; Mismatches 11; Indels 5; Gaps 2;

QY 1 MRVSIQHCALALGLALTAICS-OGAASOPDLASRLRQALAA 44
 1 MPSTRIQCALALLAVALSVCSVSGAPS-----DAKLRQFLORSILA 41

RESULT 4

150798
 preprosomatostatin SS-14 - channel catfish
 C:Species: Ictalurus punctatus (channel catfish)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
 C:Accession: I50798
 R:Dixon, J.E.; Andrews, P.C.
 Adv. Exp. Med. Biol. 188, 19-29, 1985
 A:Title: Somatostatins of the channel catfish.
 A:Reference number: I50798; MUID:85303576; PMID:2863931
 A:Accession: I50798
 A:Status: preliminary; translated from GB/EMBL/DBDJB
 A:Molecule type: mRNA
 A:Residues: 1-114 <DIK>
 A:Cross-references: GB:IM25903; NID:9213339; PIDN:AAA49339.1; PID:9213340
 C:Superfamily: somatostatin

Query Match 17.9%; Score 80.5; DB 2; Length 114;

Best Local Similarity 46.7%; Pred. No. 0.078;
 Matches 21; Conservative 8; Mismatches 11; Indels 5; Gaps 2;
 QY 1 MRVSIQHCALALGLALTAICS-OGAASOPDLASRLRQALAA 44
 Db 1 MPSTRIQCALALLAVALSVCSVSGAPS-----DAKLRQFLORSILA 41

RESULT 5

RIHDSI
 somatostatin I precursor - human
 N:Alternate names: preprosomatostatin
 C:Species: Homo sapiens (man)
 C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 18-Jun-1999
 C:Accession: A43614; A01430; S09381; S50024
 R:Shen, L.P.; Rutter, W.J.
 Science 224, 168-171, 1984
 A:Title: Sequence of the human somatostatin I gene.
 A:Reference number: A43614; MUID:84146798; PMID:6142531
 A:Accession: A43614
 A:Molecule type: DNA
 A:Residues: 1-116 <SH2>
 A:Cross-references: GB:J00306; NID:9338287; PIDN:AAA60566.1; PID:9338288
 R:Shen, L.P.; Pictet, R.L.; Rutter, W.J.
 Proc. Natl. Acad. Sci. U.S.A. 79, 4575-4579, 1982
 A:Title: Human somatostatin I: sequence of the cDNA.
 A:Reference number: A01430; MUID:83014931; PMID:6126875
 A:Accession: A01430
 A:Molecule type: mRNA
 A:Residues: 1-116 <SH2>
 A:Cross-references: GB:J00306; NID:9338287; PIDN:AAA60566.1; PID:9338288
 A:Experimental source: pancreatic somatostatinoma
 R:Goetz, S.; Boileau, G.; Zollinger, L.; Nault, C.; Rholam, M.; Cohen, P.
 EMBO J. 8, 2911-2916, 1989
 A:Title: Site-specific mutagenesis identifies amino acid residues critical in prohorm
 A:Reference number: S09381; MUID:90059875; PMID:2573512
 A:Accession: S09381
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 88-108 <COM>
 R:Odum, L.; Johnsen, A.H.
 Biochem. J. 303, 263-268, 1994
 A:Title: Human seminal plasma contains somatostatin-64.
 A:Reference number: S50024; MUID:95031969; PMID:7945250
 A:Accession: S50024
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 53-62; 67-82 <ODU>
 C:Comment: Somatostatin inhibits the release of somatotropin.
 C:Genetics:
 A:Gene: GDB:SST
 A:Cross-references: GDB:I19604; OMIM:182450
 A:Map position: 3q28-3q28
 A:Introns: 46/3
 C:Function:
 A:Description: inhibits the secretion of a number of peptide hormones, including soma

Query Match 17.7%; Score 79.5; DB 1; Length 116;
 Best Local Similarity 34.9%; Pred. No. 0.1;
 Matches 29; Conservative 13; Mismatches 28; Indels 13; Gaps 5;

QY 5 OHCALALGLALAT-CSQGAASOPDLASRLRQALAAALPHRSVSEWTFYPNC 63
 Db 5 RLOCALALLSIALVGTWVSAPDRL-----ROFLQKSLAAA-----AKQELAKYFLAE- 55

OY 64 PCLWRPRKVGKPOLKAKEDLER 86
 DB 56 --LLSEPNOTENDALE-PEDLSQ 75

RESULT 6

A28968

soomatostatin I precursor - crab-eating macaque
 N:Alternate names: preprosomatostatin
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
 C:Species: Macaca fascicularis (crab-eating macaque)
 C:Date: 30-Jun-1989 #sequence_revision 31-Jan-1997 #text_change 18-Jun-1999
 C:Accession: A28968
 R:Rivais, G.H.; Sutcliffe, J.G.
 Proc. Natl. Acad. Sci. U.S.A. 85, 1696-1700, 1988
 A:Title: Phenol emulsion-enhanced DNA-driven subtractive cDNA cloning: isolation of low
 A:Reference number: A28968; MUID:88144503; PMID:2894033
 A:Accession: A28968
 A:Molecule type: mRNA
 A:Residues: 1-116 <TRA>
 A:Cross-references: GB:M19318; NID:g342298; PIDN:AAA36908.1; PID:g342299
 A:Comment: Somatostatin inhibits the release of somatotropin.
 C:Superfamily: somatostatin
 C:Keywords: hormone; neuropeptide
 F:1-24/Domain: signal sequence #status: predicted <SIG>
 F:25-88/Domain: propeptide #status: predicted <PRO>
 F:89-116/Product: somatostatin-28 #status: predicted <M28>
 F:103-116/Product: somatostatin-14 #status: predicted <M14>
 F:105-116/Disulfide bonds: #status: predicted

Query Match 17.7%; Score 79.5; DB 1; Length 116;
 Best Local Similarity 34.9%; Pred. No. 0.1; Mismatches 13; Gaps 5;
 Matches 29; Conservative 13; Indels 28; Indels 13; Gaps 5;
 OY 5 QIHCALALLGLALAI-CSOGASOPDLASRLRLQRLAALPHRSVGVSRWTFYPMC 63
 DB 5 RIQCALALSIVLALGVGAFSDPRL----RFLQKSLAA-----AGKQLAKYFLAP- 55

OY 64 PCLWRPRKVGKPOLKAKEDLER 86
 DB 56 --LLSEPNOTENDALE-PEDLSQ 75

RESULT 7

JC6166

soomatostatin-14 precursor - laughing frog
 N:Alternate names: PSS1 protein
 C:Species: Rana ridibunda (laughing frog)
 C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 16-Jul-1999
 C:Accession: JC6166
 R:Tostivint, H.; Lihrmann, I.; Bucharies, C.; Vieau, D.; Coulouarn, Y.; Fournier, A.; Co
 Proc. Natl. Acad. Sci. U.S.A. 93, 12605-12610, 1996
 A:Title: Occurrence of two somatostatin variants in the frog brain: Characterization of
 A:Contents: brain
 A:Accession: JC6166
 A:Molecule type: mRNA
 A:Residues: 1-115 <TOS>
 A:Cross-references: GB:U68136; NID:g1890650; PIDN:AA60093.1; PID:g1890651
 A:Comment: This protein acts both as a neurotransmitter/neuromodulator and a hormone.
 C:Genetics:
 A:Gene: pss1
 C:Superfamily: somatostatin
 C:Keywords: brain; hormone

Query Match 17.0%; Score 76.5; DB 2; Length 115;
 Best Local Similarity 47.8%; Pred. No. 0.22;
 Matches 22; Conservative 8; Mismatches 11; Indels 5; Gaps 2;
 OY 1 MRVSOIHCALALLGLALAI-CSOGAA-SQPDLDLASRLRLQRLAALAA 45
 DB 1 MQSCRVOCATLILSLALAINISAPTPRL-----RFLQKSLAA 42

RESULT 8

RIAFS2

soomatostatin II precursor - American goosefish
 C:Species: Lophius americanus (American goosefish)
 C:Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 28-May-1999
 C:Accession: B93236; A94038; A27376; A01434; A1881; A93236
 R:Hobart, P.; Crawford, R.; Shen, L.; Picot, R.; Ruter, W.J.
 Nature 288, 137-141, 1980
 A:Title: Cloning and sequence analysis of cDNAs encoding two distinct somatostatin
 A:Reference number: A93236; MUID:81052423; PMID:6107860
 A:Accession: B93236
 A:Molecule type: mRNA
 A:Residues: 1-125 <NOB>

A:Cross-references: GB:V00641; GB:J00947; GB:M23199; NID:g64030; PIDN:CAA23987.1; P
 A:Experimental source: islet tissue (endocrine pancreas)
 R:Spless, J.; Noe, B.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 277-281, 1985
 A:Title: Processing of an anglerfish somatostatin precursor to a hydroxylysine-cont
 A:Reference number: A94038; MUID:85113184; PMID:2857489
 A:Accession: A94038
 A:Molecule type: protein
 A:Residues: 98-125 <SPI>
 R:Andrews, P.C.; Nichols, R.; Dixon, J.E.
 J. Biol. Chem. 262, 12692-12699, 1987
 A:Title: Post-translational processing of preprosomatostatin-II examined using fast
 A:Reference number: A27376; MUID:87308304; PMID:2887572
 A:Accession: A27376
 A:Molecule type: protein
 A:Residues: 1-76, 'DV', '79-89', 'G', '91-125' <AND>

C:Superfamily: somatostatin
 C:Keywords: hydroxylysine; neuropeptide; pyroglutamic acid
 F:1-24/Domain: signal sequence #status: experimental <SIG>
 F:25-97/Domain: propeptide #status: experimental <PRO>
 F:97-125/Product: somatostatin II #status: experimental <M25>
 F:25/Modified site: pyroglutamic acid (Gln) (in mature form) #status: expe
 F:114-125/Disulfide bonds: #status: experimental
 F:120/Modified site: hydroxylysine (Lys) #status: experimental

Query Match 15.4%; Score 69.5; DB 1; Length 125;
 Best Local Similarity 36.1%; Pred. No. 1.5;
 Matches 22; Conservative 7; Mismatches 17; Indels 15; Gaps 3;

OY 6 IHCALALLGLALAI-CSOGASO-----PDLASRR--LLQRLAALPHRSVGSER 55
 DB 4 IRCFALLALLALVLCGPSYSSOLDREQSDNDLDLRLQHWLLERARSAGL-----LSQE 58

OY 56 W 56
 DB 59 W 59

RESULT 9

RIBOS1

soomatostatin precursor - bovine
 N:Alternate names: preprosomatostatin
 N:Contains: somatostatin 14 (SS-14); somatostatin 28 (SS-28)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 06-Mar-1992 #sequence_revision 31-Jan-1997 #text_change 18-Jun-1999
 C:Accession: A40929
 R:Su, C.J.; White, J.W.; Li, W.H.; Luo, C.C.; Frazier, M.L.; Saunders, G.F.; Chan,
 Mol. Endocrinol. 2, 209-216, 1988
 A:Title: Structure and evolution of somatostatin genes.
 A:Reference number: A40929; MUID:88288237; PMID:2899837
 A:Accession: A40929
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-116 <SUA>
 A:Cross-references: GB:M31217; NID:g163636; PIDN:AAA30744.1; PID:g163637
 A:Note: the authors translated the codon ATT for residue 65 as Asn
 C:Comment: Somatostatin inhibits the release of somatotropin.
 C:Superfamily: somatostatin
 C:Keywords: hormone; neuropeptide

A:Accession: A46033
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-860, 1-862-1205 <I&M2>
A:Cross-references: GB:M89952; NID:g162976
A:Experimental source: endothelial
A:Note: sequence extracted from NCBI backbone (NCBIP:108720)
R:Nishida, K.; Harrison, D.G.; Navas, J.P.; Fisher, A.A.; Dockery, S.P.; Nerem, R.M.; Al
J. Clin. Invest. 90, 2092-2096, 1992
A:Title: Molecular Cloning and Characterization of the constitutive bovine aortic Endoth
A:Reference number: I45945; MUID:93055452; PMID:1385480
A:Accession: I45945
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1205 <NIS>
A:Cross-references: GB:M99057; NID:g163421; PIDN:AAA30667.1; PID:g163422
R:Sessa, W.C.; Harrison, J.K.; Barber, C.M.; Zeng, D.; Durieux, M.E.; D'Angelo, D.D.; Ly
J. Biol. Chem. 267, 15274-15276, 1992
A:Title: Molecular cloning and expression of a cDNA encoding endothelial cell nitric ox
A:Reference number: A42841; MUID:92348367; PMID:1379225
A:Accession: A42841
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-551, N', 53-99, R', 101-147, M', 149-164, T', 166-317, GA', 320, HTGVVRGP', 329-3
K', 517-632, G', 694-740, A', 742-753, N', 755-799, N', 801-803, SA', 806-856, V', 858-906, LV
A:Experimental source: aortic endothelial cells
A:Note: sequence extracted from NCBI backbone (NCBIP:109564); contains a number of typog
A:Accession: I45946
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-99, R', 101-164, T', 166-317, GA', 320, HTGVVRGP', 329-454, Y', 456-458, P', 460
A:Cross-references: GB:M95674; NID:g163426; PIDN:AAA30669.1; PID:g163427
A:Experimental source: aortic endothelial cells
A:Note: submitted to GenBank, August 1992
A:Note: GenBank entry BOVNOS, release 103.0, has a typographical error in the reference
C:Function:
A:Description: catalyzes the oxidation of an L-arginine quinidino nitrogen and of NADPH
C:Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferrihemoprotein reduct
C:Keywords: blocked amino end; calmodulin binding; chromoprotein; FAD; flavoprotein; FMN
F:493-512/Region: calmodulin binding #status predicted
F:532-1161/Domain: NADPH-ferrihemoprotein reductase homology <FEH>
F:524-705/Domain: flavodoxin homology <FLX>
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:3/Modified site: aspartic acid (Asn) #status predicted
F:106/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 14.98; Score 67; DB 1; Length 1205;
Best Local Similarity 35.58; Pred. No. 25;
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;
QY 8 CAAALGLALATC-SQGAAS---OPDLASRRLLQRLAALPHRSGVSRWRTFFPNC 63
DB 15 CGUG-LGLGLGCGKQSPAPPEP-----SRAPAPATPHAPDHSPA-----PNS 58
QY 64 PCLRWRPRKVGPKQLK 79
DB 59 PTLTRPPEGPKPRVK 74
RESULT 14
H81009
exodeoxyribonuclease NMB2082 [imported] - Neisseria meningitidis (strain MC58 serogroup
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
C:Accession: H81009; A82031
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qian, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: H81009

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <TET>
A:Cross-references: GB:AE002557; GB:AE002098; NID:g7227332; PIDN:AAF42400.1; PID:g7
A:Experimental source: serogroup B, strain MC58
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.;
Hollroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajanc
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 224
A:Reference number: A81775; MUID:2022556; PMID:10761919
A:Accession: A82031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <PAR>
A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83652.1; PID:g7
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMB2082; NMA0348
C:Superfamily: exodeoxyribonuclease III
Query Match 14.88; Score 66.5; DB 2; Length 259;
Best Local Similarity 24.68; Pred. No. 6.4; Mismatches 13; Indels 11; Gaps 2;
Matches 16; Conservative 13; Mismatches 25; Indels 11; Gaps 2;
QY 14 GLAALCSQGAASOPDLASRRLLQRLAALP-----HRSGVSRWRTFFPNC 63
DB 141 GRDIVVCGDNIAHQHIDLNKMGKN-SGPLPEEREWIGKVIHKLGTDMWRTLYPDV 199
QY 64 PCLRW 68
DB 200 PGYTW 204
RESULT 15
T43456
hypothetical protein DKFZp434L061.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Mar-2000
C:Accession: T43456
R:Postka, A.; Klein, M.; Meves, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: 222516
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-571 <AAA>
A:Cross-references: EMBL:AL133647
A:Experimental source: adult testis; clone DKFZp434L061
C:Genetics:
A:Note: DKFZp434L061.1
C:Superfamily: hydroxyproline-rich glycoprotein
Query Match 14.88; Score 66.5; DB 2; Length 571;
Best Local Similarity 29.08; Pred. No. 14;
Matches 20; Conservative 15; Mismatches 19; Indels 15; Gaps 4;
QY 1 MRVSOIHCAALIGLALATC-SQGAASQPDLDASRRLLQRLAALP-----LPHRSGVSRW 56
DB 405 LQLSEVN-----IPLSLGVCPLG-----PVPFLTKELYQQAAEAAHHHPSPD-SERI 453
QY 57 RTEVPNCPC 65
DB 454 RQYLPRNPC 462
Search completed: March 21, 2003, 11:41:40
Job time : 20.4549 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 16.6094 Seconds
(without alignments)
276.816 Million cell updates/sec

Title: US-09-727-739b-17

Perfect score: 450

Sequence: 1 MRVSIHICALLGLALAIC.....RWRPRKVRGPKAKEDLER 86

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 221153 seqs, 53462247 residues 221153

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/1/pubpa/US05_NEW_PUB_PEP.*
- 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB_PEP.*
- 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB_PEP.*
- 6: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB_PEP.*
- 7: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB_PEP.*
- 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB_PEP.*
- 9: /cgn2_6/ptodata/1/pubpa/US09_NEW_PUB_PEP.*
- 10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB_PEP.*
- 11: /cgn2_6/ptodata/1/pubpa/US10_NEW_PUB_PEP.*
- 12: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB_PEP.*
- 13: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB_PEP.*
- 14: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67.5	15.0	95	10	US-09-864-761-41361
2	65.5	14.6	513	10	Sequence 41361, A
3	63.5	14.1	333	9	Sequence 63, Appl
4	62.5	13.9	110	10	Sequence 131, Appl
5	62.5	13.9	110	12	Sequence 3, Appl
6	62	13.8	833	9	Sequence 54, Appl
7	61.5	13.7	333	9	Sequence 42, Appl
8	61.5	13.7	333	9	Sequence 132, Appl
9	61.5	13.7	333	9	Sequence 132, Appl
10	61.5	13.7	333	9	Sequence 132, Appl
11	61.5	13.7	333	9	Sequence 132, Appl
12	61.5	13.7	333	9	Sequence 132, Appl
13	61.5	13.7	333	9	Sequence 132, Appl
14	61.5	13.7	333	9	Sequence 132, Appl
15	61.5	13.7	333	9	Sequence 132, Appl
16	61.5	13.7	333	9	Sequence 132, Appl
17	61.5	13.7	333	9	Sequence 132, Appl
18	61.5	13.7	333	9	Sequence 132, Appl
19	61.5	13.7	333	9	Sequence 132, Appl

20	61.5	13.7	333	9	US-10-142-419-132	Sequence 132, App
21	61.5	13.7	333	9	US-10-123-262-132	Sequence 132, App
22	61.5	13.7	333	9	US-10-142-423-132	Sequence 132, App
23	61	13.6	697	10	US-09-770-643A-18	Sequence 18, Appl
24	61	13.6	745	10	US-09-770-643A-16	Sequence 16, Appl
25	61	13.6	791	10	US-09-770-643A-22	Sequence 22, Appl
26	61	13.6	839	10	US-09-770-643A-20	Sequence 20, Appl
27	61	13.6	1175	10	US-09-770-643A-26	Sequence 26, Appl
28	61	13.6	1175	10	US-09-770-643A-30	Sequence 30, Appl
29	61	13.6	1259	10	US-09-770-643A-4	Sequence 4, Appl
30	61	13.6	1298	10	US-09-770-643A-28	Sequence 28, Appl
31	61	13.6	1307	10	US-09-770-643A-2	Sequence 2, Appl
32	61	13.6	1307	10	US-09-789-561-128	Sequence 128, App
33	60.5	13.4	90	10	US-09-764-868-665	Sequence 665, App
34	59	13.1	468	9	US-08-900-220C-17	Sequence 17, Appl
35	58.5	13.0	396	8	US-09-883-848A-17	Sequence 17, Appl
36	58.5	13.0	396	10	US-09-244-466-2	Sequence 2, Appl
37	58.5	13.0	396	10	US-09-151-999-17	Sequence 17, Appl
38	58.5	13.0	396	10	US-09-934-392-6	Sequence 6, Appl
39	58	12.9	234	9	US-09-739-907-171	Sequence 171, App
40	58	12.9	274	10	US-09-739-907-79	Sequence 79, Appl
41	58	12.9	525	10	US-10-007-706-1	Sequence 1, Appl
42	58	12.9	1503	9	US-09-815-242-5184	Sequence 5184, Ap
43	57.5	12.8	384	10	US-09-470-276-2	Sequence 2, Appl
44	57	12.7	834	9	US-09-934-909-29	Sequence 29, Appl
45	57	12.7	834	10		

ALIGNMENTS

RESULT 1
US-09-864-761-41361
; Sequence 41361 Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, David R.
; APPLICANT: Rank, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aemlica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41361
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121914.16
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
; -09-864-761-41361
Query Match 15.0%; Score 67.5; DB 10; Length 95;
Best Local Similarity 27.7%; Pred. No. 0.81; Mismatches 23; Conservative 11; Gaps 4;
Matches 23; Conservative 11; Mismatches 29; Indels 29; Gaps 4;
QY 10 LALLGLALAI-----CSOGAASQPDLDASRRLLQRLAALPHRSGVSE 54
Db 2 LALLLSRVNLLRGHPAYLNAGTSLGRSKPSAHFTA-----ASGAASPRSSIRA 55
QY 55 RW-----RTFFPNCPLRWPRK 72
Db 56 SWSWMDRSWWTTC-----AWRPRR 75
RESULT 2
US-09-833-745-63
; Sequence 63, Application US/09833745
; Patent No. US20020052038A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 53
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Rhizobium meliloti
US-09-833-745-63
Query Match 14.6%; Score 65.5; DB 10; Length 513;
Best Local Similarity 38.2%; Pred. No. 9.5;
Matches 21; Conservative 5; Mismatches 24; Indels 5; Gaps 1;
QY 4 SQIHCALLGLAL-----AICSGAASQPDLDASRRLLQRLAALPHRSGVS 53
Db 330 SQAHVGAALGLANDSLAVAIEVAIAISERRIDRLVNLPSRGLPAFLAGDSGVS 384
RESULT 3
US-09-984-271-131
; Sequence 131, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 131
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-984-271-131
Query Match 14.1%; Score 63.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 9.9;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;
QY 18 AICSQ-GRASQP-DLDASRRLLQRLAALPHRSGVSEWRFTFY---PNCPLRW 69
Db 72 AVCGQGPMPDPMDLPVGQRRTLLRXIL-----VSDRYRFLCYVPKVACSNWK 120
RESULT 4
US-09-766-396-3
; Sequence 3, Application US/09766396
; Patent No. US20020013456A1
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; de Lecea, Luis
; Siggins, George R.
; Henriksen, Steven J.
; TITLE OF INVENTION: CORPUSCULUM, NEUROPEPTIDES,
; COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 No. US20020013456A1th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,396
; FILING DATE: 18-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/857,389
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Schmonsees, William
; REGISTRATION NUMBER: 31,796
```

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: C-terminal

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-766-396-3

Query Match 13.9% Score 62.5; DB 10; Length 110;
Best Local Similarity 48.7% Pred. No. 3.7;
Matches 19; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 8 CALALGLALATCS-QGAASQPDLDLASRRLQRLALAAA 45

Db 2 CALAALCIVLALGGVTGAPDRL-----ROFLQKSLAAA 36

RESULT 5

US-10-062-375-3

Sequence 3, Application US/10062375

Patent No. US20020133000A1

GENERAL INFORMATION:

APPLICANT: Sutcliffe, Gregor J.

de Lecea, Luis

Siggins, George R.

Henriksen, Steven J.

TITLE OF INVENTION: CORTISTATIN: NEUROPEPTIDES,

COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESS: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 NO. US20020133000A1th Torrey Pines Road, TPC-8

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/062,375

FILING DATE: 30-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/857,389

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schmonsees, William

REGISTRATION NUMBER: 31,796

REFERENCE/DOCKET NUMBER: 22908-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-7041

TELEFAX: (415) 324-0638

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 110 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: C-terminal

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-062-375-3

Query Match 13.9% Score 62.5; DB 12; Length 110;

Best Local Similarity 48.7% Pred. No. 3.7;

Matches 19; Conservative 5; Mismatches 10; Indels 5; Gaps 2;

QY 8 CALALGLALATCS-QGAASQPDLDLASRRLQRLALAAA 45

Db 2 CALAALCIVLALGGVTGAPDRL-----ROFLQKSLAAA 36

RESULT 6

US-09-470-276-54

Sequence 54, Application US/09470276

Publication No. US20020187469A1

GENERAL INFORMATION:

APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.

APPLICANT: KLODNER, Richard

APPLICANT: WINAND, Neda

TITLE OF INVENTION: A METHOD OF DETECTION OF ALTERATIONS IN MSH5

FILE REFERENCE: 700157/47483C US/09/470,276

CURRENT APPLICATION NUMBER: US/09/470,276

CURRENT FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: 60/051,686

PRIOR FILING DATE: 1997-07-03

PRIOR APPLICATION NUMBER: PCT/US98/13850

PRIOR FILING DATE: 1998-07-02

NUMBER OF SEQ ID NOS: 104

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 54

LENGTH: 833

TYPE: PRT

ORGANISM: Human

US-09-470-276-54

Query Match 13.8% Score 62; DB 9; Length 833;

Best Local Similarity 27.5% Pred. No. 42;

Matches 30; Conservative 12; Mismatches 37; Indels 30; Gaps 5;

QY 3 VSQIHC-----ALALGLALATCSQGAASQPDLDLASRRLQRLALAAAAL-----PHR 49

Db 489 LGLDCEIRDOETLMTQLCQVLRASVLTIVLRLDLASRLDVLALASAAARDYGYGRPHY 548

QY 50 S-----GVSE-----WRTFYFN---CPCLMRPRKVKGPOLAK 81

Db 549 SPCHGVIRNRHPLMELCARTFVFNSTDCGQDGRVKVITGPNSSGK 597

RESULT 7

US-10-001-054-42

Sequence 42, Application US/10001054

Publication No. US20020192209A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Baker, Kevin

APPLICANT: Goddard, Audrey

APPLICANT: Gurney, Austin

APPLICANT: Hebert, Carolyn

APPLICANT: Henzel, William

APPLICANT: Kabakoff, Rhona

APPLICANT: Shelton, David

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin

APPLICANT: Wood, William

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC

CELL GROWTH

FILE REFERENCE: P3034R1PCT

CURRENT APPLICATION NUMBER: US/10/001,054

CURRENT FILING DATE: 2001-11-30

PRIOR APPLICATION NUMBER: 60/059114

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079920

PRIOR FILING DATE: 1998-03-30

PRIOR APPLICATION NUMBER: 60/082999

PRIOR FILING DATE: 1998-04-24

PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/096891
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/096894
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100263
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/107783
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112420
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116533
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/131294
PRIOR FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/218517
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 09/284291
PRIOR FILING DATE: 1999-04-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380913
PRIOR FILING DATE: 1999-09-09
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/866034
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/882636
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US99/00106
PRIOR FILING DATE: 1999-01-05
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/08615
PRIOR FILING DATE: 1999-04-20
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00376
PRIOR FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04341
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/06884
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/13705
PRIOR FILING DATE: 2000-05-17
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/22031
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/30873
PRIOR FILING DATE: 2000-11-10

PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
Query Match 13.7%; Score 61.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 17;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;
QY 18 AICSO-GAASOP-DLDLASRLLORALAAALPHRSGVSRWTFY---PNCPCLRWR 69
DB 72 AVCQGMFPRDPLVQQRRTLLRHIL-----VSDRFLCYVYKVCASNWK 120

RESULT 9

US-10-121-049-132
Sequence 132, Application US/10121049
Publication No. US2003002239A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zeng, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 132
LENGTH: 333
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-132

Query Match 13.7%; Score 61.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 17;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;
QY 18 AICSO-GAASOP-DLDLASRLLORALAAALPHRSGVSRWTFY---PNCPCLRWR 69
DB 72 AVCQGMFPRDPLVQQRRTLLRHIL-----VSDRFLCYVYKVCASNWK 120

RESULT 10
US-10-123-904-132

```

; Sequence 132, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 132
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-132

```

```

Query Match 13.7%; Score 61.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 17; Indels 13; Gaps 4;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;

OY 18 AICSO-GAASOP-DLDLASHRLLQRLAALPHRSGVSRWRTY---PNCPCLRWR 69
Db 72 AVCGPGMPDPMDLPVQGRLLRHIL-----VSDRYRFLYCYVPKVCASNK 120

```

```

RESULT 11
US-10-140-470-132
; Sequence 132, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 132
; LENGTH: 333

```

```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-132

```

```

Query Match 13.7%; Score 61.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 17; Indels 13; Gaps 4;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;

OY 18 AICSO-GAASOP-DLDLASHRLLQRLAALPHRSGVSRWRTY---PNCPCLRWR 69
Db 72 AVCGPGMPDPMDLPVQGRLLRHIL-----VSDRYRFLYCYVPKVCASNK 120

```

```

RESULT 12
US-10-175-746-132
; Sequence 132, Application US/10175746
; Publication No. US20030022770A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 132
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-132

```

```

Query Match 13.7%; Score 61.5; DB 9; Length 333;
Best Local Similarity 33.3%; Pred. No. 17; Indels 13; Gaps 4;
Matches 19; Conservative 6; Mismatches 19; Indels 13; Gaps 4;

OY 18 AICSO-GAASOP-DLDLASHRLLQRLAALPHRSGVSRWRTY---PNCPCLRWR 69
Db 72 AVCGPGMPDPMDLPVQGRLLRHIL-----VSDRYRFLYCYVPKVCASNK 120

```

```

RESULT 13
US-10-176-918-132
; Sequence 132, Application US/10176918
; Publication No. US20030022775A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:41:02 ; Search time 16.2403 Seconds
(without alignments)
155.808 Million cell updates/sec

Title: US-09-727-739B-17
Perfect score: 450
Sequence: 1 MRVSIHICALALLGLAIC.....RWRPKYKGPOLKAKEDLER 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

-- Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	14.9	1205	1	US-07-908-245-2
2	67	14.9	1205	2	US-08-319-868-10
3	67	14.9	1205	4	US-09-123-708-6
4	67	14.9	1205	4	US-09-123-708-6
5	62.5	13.9	1110	3	US-08-548-322-3
6	59.5	13.2	2476	2	US-08-276-967-2
7	58.5	13.0	396	4	US-09-325-256-24
8	58	12.9	802	4	US-09-632-098-2
9	57	12.7	479	1	US-08-484-105-10
10	57	12.7	479	1	US-08-484-105-10
11	57	12.7	834	4	US-09-143-571-29
12	56.5	12.6	47	4	US-08-469-260A-344
13	56.5	12.6	396	1	US-08-176-427B-4
14	56.5	12.6	396	2	US-08-356-060A-9
15	56.5	12.6	396	4	US-08-460-900C-9
16	56.5	12.6	396	4	US-08-674-509B-9
17	56.5	12.6	396	4	US-08-954-698-9
18	56.5	12.6	396	4	US-09-293-505-13
19	56.5	12.6	396	4	US-08-957-874-9
20	56.5	12.6	396	4	US-09-325-256-18
21	55.5	12.3	1296	3	US-08-728-603-15
22	54.5	12.1	741	2	US-08-462-481-2
23	54.5	12.1	741	2	US-08-436-771-2
24	54.5	12.1	741	2	US-08-436-771-4
25	54.5	12.1	741	2	US-08-434-998-2
26	54.5	12.1	741	2	US-08-434-998-4
27	54.5	12.1	741	2	US-08-487-797-2

28	54.5	12.1	741	2	US-08-487-797-4	Sequence 4, Appli
29	54.5	12.1	741	2	US-08-701-005A-2	Sequence 2, Appli
30	54.5	12.1	741	2	US-08-479-895-2	Sequence 2, Appli
31	54.5	12.1	741	3	US-08-943-956A-2	Sequence 2, Appli
32	54.5	12.1	741	5	PCT-US95-02058-2	Sequence 2, Appli
33	54.5	12.1	741	5	PCT-US95-02058-4	Sequence 4, Appli
34	54	12.0	315	4	US-09-370-398-6	Sequence 6, Appli
35	54	12.0	374	1	US-08-468-847B-12	Sequence 12, Appli
36	54	12.0	375	2	US-08-459-101A-2	Sequence 2, Appli
37	54	12.0	408	3	US-08-458-434A-2	Sequence 2, Appli
38	54	12.0	671	4	US-09-132-118-2	Sequence 2, Appli
39	54	12.0	709	1	US-08-444-005-17	Sequence 17, Appli
40	53.5	11.9	112	4	US-09-344-587-14	Sequence 14, Appli
41	53.5	11.9	120	1	US-08-347-492B-2	Sequence 2, Appli
42	53.5	11.9	120	2	US-08-798-143-2	Sequence 2, Appli
43	53.5	11.9	120	5	PCT-US95-15484-2	Sequence 2, Appli
44	53.5	11.9	136	5	PCT-US95-07171-2	Sequence 2, Appli
45	53.5	11.9	191	5	PCT-US95-07171-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-07-908-245-2
; Sequence 2, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T.J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta
; CELL TYPE: Endothelial
; FEATURE:
; NAME/KEY: Binding-site

```

LOCATION: 496..512
OTHER INFORMATION: /note= "CA++/CAM binding domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 651..678
OTHER INFORMATION: /note= "FMN binding domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 795..806
OTHER INFORMATION: /note= "FAD-Pyrophosphate binding"
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 937..947
OTHER INFORMATION: /note= "FAD-Isolalloxanthine"
OTHER INFORMATION: binding domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 1012..1030
OTHER INFORMATION: /note= "NADPH-Ribose binding"
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 1111..1124
OTHER INFORMATION: /note= "NADPH-Ribose binding"
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 33..34
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 46..47
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 53..54
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 58..59
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 97..98
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 116..117
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 282..283
OTHER INFORMATION: /note= "Potential proline"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 459..460
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 472..473
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain

```

```

LOCATION: 602..603
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 727..728
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 838..839
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 869..870
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 872..873
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1085..1086
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1202..1203
OTHER INFORMATION: /note= "Potential proline directed"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 114..116
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 141..143
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 168..170
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 633..635
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 836..838
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1051..1053
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 738..740
OTHER INFORMATION: /note= "cAMP dependent"
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain

```

```

US-07-908-245-2
Query Match 14.9% Score 67; DB 1; Length 1205;
Best Local Similarity 35.5%; Pred. No. 7;
Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;

```

QY 8 CALALLGLALAIC-SOGAAS---OPDLASRRLLQALAAALPHRSGVSRWRTYPNC 63
 Db 15 CGLG-LGLGLGCKGQGPASPAEP-----SRAPATPHADHSPA-----PNS 58
 QY 64 PCLWRPRKVKGPOLK 79
 Db 59 PTLTRPPEGPFRVK 74

RESULT 2

US-08-319-866-10

; Sequence 10, Application US/08319866

; Patent No. 5929223

; GENERAL INFORMATION:

; APPLICANT: Tully, Timothy P.

; APPLICANT: Yin, Jerry C.

; APPLICANT: Regulski, Michael

; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES

; TITLE OF INVENTION: ASSOCIATED WITH LONG-TERM MEMORY

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/319,866

; FILING DATE: 7-OCT-1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: CSHL94-03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 861-6240

; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1205 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-319-866-10

Query Match 14.9%; Score 67; DB 2; Length 1205;

Best Local Similarity 35.5%; Pred. No. 7;

Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;

QY 8 CALALLGLALAIC-SOGAAS---OPDLASRRLLQALAAALPHRSGVSRWRTYPNC 63

Db 15 CGLG-LGLGLGCKGQGPASPAEP-----SRAPATPHADHSPA-----PNS 58

QY 64 PCLWRPRKVKGPOLK 79

Db 59 PTLTRPPEGPFRVK 74

RESULT 3

US-09-123-708-6

; Sequence 6, Application US/09123708

; Patent No. 6146887

GENERAL INFORMATION:
 APPLICANT: SCHRADER, Jurgen
 APPLICANT: CODECKE, Axel
 TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
 TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
 FILE REFERENCE: 511169-2003
 CURRENT APPLICATION NUMBER: US/09/123,708
 CURRENT FILING DATE: 1998-07-28
 EARLIER APPLICATION NUMBER: 08/553,503
 EARLIER FILING DATE: 1996-03-01
 EARLIER APPLICATION NUMBER: P4411402.8
 EARLIER FILING DATE: 1994-03-31
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Ver..2.0
 SEQ ID NO 6
 LENGTH: 1205
 TYPE: PRT
 ORGANISM: Cytomegalovirus
 US-09-123-708-6

Query Match 14.9%; Score 67; DB 4; Length 1205;

Best Local Similarity 35.5%; Pred. No. 7;

Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;

QY 8 CALALLGLALAIC-SOGAAS---OPDLASRRLLQALAAALPHRSGVSRWRTYPNC 63

Db 15 CGLG-LGLGLGCKGQGPASPAEP-----SRAPATPHADHSPA-----PNS 58

QY 64 PCLWRPRKVKGPOLK 79

Db 59 PTLTRPPEGPFRVK 74

RESULT 4

US-09-123-624-6

; Sequence 6, Application US/09123624

; Patent No. 6149936

; GENERAL INFORMATION:

; APPLICANT: SCHRADER, Jurgen

; APPLICANT: CODECKE, Axel

; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC

; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS

; FILE REFERENCE: 511169-2004

; CURRENT APPLICATION NUMBER: US/09/123,624

; CURRENT FILING DATE: 1998-07-28

; PRIOR APPLICATION NUMBER: 08/553,503

; PRIOR FILING DATE: 1996-03-01

; PRIOR APPLICATION NUMBER: 4411402.8

; PRIOR FILING DATE: 1994-03-31

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 1205

; TYPE: PRT

; ORGANISM: Bos taurus

US-09-123-624-6

Query Match 14.9%; Score 67; DB 4; Length 1205;

Best Local Similarity 35.5%; Pred. No. 7;

Matches 27; Conservative 4; Mismatches 25; Indels 20; Gaps 5;

QY 8 CALALLGLALAIC-SOGAAS---OPDLASRRLLQALAAALPHRSGVSRWRTYPNC 63

Db 15 CGLG-LGLGLGCKGQGPASPAEP-----SRAPATPHADHSPA-----PNS 58

QY 64 PCLWRPRKVKGPOLK 79

Db 59 PTLTRPPEGPFRVK 74

RESULT 5

US-08-648-322-3

; Sequence 3, Application US/08648322

DB 11 CCLALLALPAQSCGPRGVPVGRRRYARKOLVPLLYKQFVPGVPGVPTILGASGPAEGRVARG 70

QY 53 SERWTFYPN 62

DB 71 SERFDLVPN 80

RESULT 8

US-09-632-098-2

Sequence 2, Application US/096322098

Patent No. 6420154

GENERAL INFORMATION:

APPLICANT: Sheppard, Paul O.

APPLICANT: Baidur, Nand

FILE REFERENCE: 99-39

TITLE OF INVENTION: MAMMALIAN ADHESION PROTEASE PEPTIDES

CURRENT APPLICATION NUMBER: US/09/632,098

CURRENT FILING DATE: 2000-08-02

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 802

TYPE: PRT

ORGANISM: Homo sapiens

US-09-632-098-2

Query Match 12.9%; Score 58; DB 4; Length 802;

Best Local Similarity 30.0%; Pred. No. 58;

Matches 24; Conservative 5; Mismatches 25; Indels 26; Gaps 4;

QY 10 LALGLALALCSQGAASOPDLASRRL-----LORALAAALPHRSGVSE----- 54

DB 617 LDLLGLGLV--EPQTGCPRMVCSRRCKRNAFQELQRLTACHSHGAGLHPSTSOALV 674

QY 55 -----RWTFYPNC-PC 65

DB 675 AAWTVALCRKATWTFSCWPC 694

RESULT 9

US-08-484-105-10

Sequence 10, Application US/08484105

Patent No. 5369341

GENERAL INFORMATION:

APPLICANT: STILLMAN, Bruce

APPLICANT: BELL, Stephen P

APPLICANT: KOBAYASHI, Ryuji

APPLICANT: RINE, Jasper

APPLICANT: FOSS, Margit

APPLICANT: MCNALLY, Francis J

APPLICANT: LAURENSEN, Patricia

APPLICANT: HERSKOWITZ, Ira

APPLICANT: LI, Joachim J

APPLICANT: GAVIN, Kimberly

TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM: disk

MEDIUM TYPE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: GAVIN, Kimberly

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman Ph.D., Richard Aron

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-484-105-10

Query Match 12.7%; Score 57; DB 1; Length 479;

Best Local Similarity 30.8%; Pred. No. 41;

Matches 16; Conservative 10; Mismatches 20; Indels 6; Gaps 1;

QY 30 LDLASRRLQRLAALPHRSGVSEWRWTFYPNCPLRWRPRKVGKPOLKAK 81

DB 65 VELVSWKPLQAIART-----VOYKLTLYPNITDYDPLQVEEPFLVK 110

RESULT 10

US-08-484-106-10

Sequence 10, Application US/08484106

Patent No. 5614618

GENERAL INFORMATION:

APPLICANT: STILLMAN, Bruce

APPLICANT: BELL, Stephen P

APPLICANT: KOBAYASHI, Ryuji

APPLICANT: RINE, Jasper

APPLICANT: FOSS, Margit

APPLICANT: MCNALLY, Francis J

APPLICANT: LAURENSEN, Patricia

APPLICANT: HERSKOWITZ, Ira

APPLICANT: LI, Joachim J

APPLICANT: GAVIN, Kimberly

TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM: disk

MEDIUM TYPE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: GAVIN, Kimberly

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman Ph.D., Richard Aron

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 479 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-484-106-10

Query Match 12.7% Score 57; DB 1; Length 479;
Best Local Similarity 30.8% Pred. No. 41;
Matches 16; Conservative 10; Mismatches 20; Indels 6; Gaps 1;
QY 30 LELASRLRLQALAAALPHRSGVSRWRTEYPCNCLRWPRKVKGPOLKAK 81
DB 65 VELVSKPQLLOAIART-----VQYKLTLYPNITPTDYDPLQVEPELLVK 110

RESULT 11
US-09-143-571-29
Sequence 29, Application US/09143571
Patent No. 6333153
GENERAL INFORMATION:
APPLICANT: FISHEL, Richard
APPLICANT: GRADIA, Scott
APPLICANT: ACHARYA, Samir
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR EFFECTING ADENINE
TITLE OF INVENTION: NUCLEOTIDE MODULATION OF DNA MISMATCH RECOGNITION
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: 9855-601
CURRENT APPLICATION NUMBER: US/09/143,571
CURRENT FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: 60/093,935
EARLIER FILING DATE: 1998-07-23
EARLIER APPLICATION NUMBER: 60/066,977
EARLIER FILING DATE: 1997-11-28
EARLIER APPLICATION NUMBER: 60/057,136
EARLIER FILING DATE: 1997-08-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 29
LENGTH: 834
TYPE: PRT
ORGANISM: Homo sapiens
US-09-143-571-29

Query Match 12.7% Score 57; DB 4; Length 834;
Best Local Similarity 25.7% Pred. No. 82;
Matches 28; Conservative 12; Mismatches 39; Indels 30; Gaps 4;
QY 3 VSQIHC-----ALALGLALATCSOGAASQPDLDLASRLRLQALAAAL-----46
DB 490 LGDLCEIRDOETLMTQCCVLAARAVLTRVLDLASRLDVLALASAAADYGRPRY 549
QY 47 -PHRSGVSR-----WRTFTPN---CPCLRWPRKVKGPOLKAK 81
DB 550 SPQVLGVRIQNGRHPMLCARTFVNSTECGGDKGRVKVITGPNSSGK 598

RESULT 12
US-08-469-260A-344
Sequence 344, Application US/08469260A
Patent No. 6451578
GENERAL INFORMATION:
APPLICANT: JOHN N. SIMONS
APPLICANT: TAMI J. PILOT-MATIAS
APPLICANT: GEORGE J. DAWSON
APPLICANT: GEORGE G. SCHLAUDER
APPLICANT: SURESH M. DESAI
APPLICANT: THOMAS P. LEARY
APPLICANT: ANTHONY SCOTT MUERHOFF
APPLICANT: JAMES C. ERKER
APPLICANT: SHERI L. BULIK
APPLICANT: ISA K. MUSHAWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/APGD

STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/469,260A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 344:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-260A-344

Query Match 12.68; Score 56.5; DB 4; Length 47;
Best Local Similarity 41.08; Pred. No. 2.5;
Matches 16; Conservative 4; Mismatches 16; Indels 3; Gaps 1;
QY 21 SQGAASQPDLDLASRLRLQAL---AAALPHRSGVSRW 56
DB 2 SEGASRPDLRWRTQLERAVYCACARLQARPGHQRW 40

RESULT 13
US-08-176-427B-4
Sequence 4, Application US/08176427B
Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:

```
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-4

Query Match 12.6%; Score 56.5; DB 1; Length 396;
Best Local Similarity 28.6%; Pred. No. 37;
Matches 20; Conservative 7; Mismatches 28; Indels 15; Gaps 2;

QY 8 CALALGLALATCSOGAASQPDLDLASRRL-----ORALAAALPHRSV--- 52
Db 11 CCLALLALSAQSCGPGVGRRRYVRKQLVPLLYKQFVSPMPERTLGASGPAEGRVTRG 70
-QY 53 SERWRTFYPN 62
Db 71 SERFDLVPN 80

RESULT 14
US-08-356-060A-9
; Sequence 9, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-356-060A-9

Query Match 12.6%; Score 56.5; DB 2; Length 396;
Best Local Similarity 28.6%; Pred. No. 37;
Matches 20; Conservative 7; Mismatches 28; Indels 15; Gaps 2;

QY 8 CALALGLALATCSOGAASQPDLDLASRRL-----ORALAAALPHRSV--- 52
```

```
Db 11 CCLALLALSAQSCGPGVGRRRYVRKQLVPLLYKQFVSPMPERTLGASGPAEGRVTRG 70
QY 53 SERWRTFYPN 62
Db 71 SERFDLVPN 80

RESULT 15
US-08-460-900C-9
; Sequence 9, Application US/08460900C
; Patent No. 6155747
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; APPLICANT: Bunciot, David A.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,900C
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 4-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006.05
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-900C-9

Query Match 12.6%; Score 56.5; DB 4; Length 396;
Best Local Similarity 28.6%; Pred. No. 37;
Matches 20; Conservative 7; Mismatches 28; Indels 15; Gaps 2;

QY 8 CALALGLALATCSOGAASQPDLDLASRRL-----ORALAAALPHRSV--- 52
Db 11 CCLALLALSAQSCGPGVGRRRYVRKQLVPLLYKQFVSPMPERTLGASGPAEGRVTRG 70
QY 53 SERWRTFYPN 62
Db 71 SERFDLVPN 80
```

Tue Mar 25 08:25:34 2003

us-09-727-739b-17.rai

Page 8

Search completed: March 21, 2003, 11:42:33
Job time : 20.2403 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:35:23 ; Search time 53.1502 Seconds
(without alignments)
215.607 Million cell updates/sec

Title: US-09-727-739b-17

Perfect score: 450

Sequence: 1 MRVSQIHCALLGLAIC.....RWRPRKVGQPKAKEDLER 86

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	450	100.0	111	22 AAU07668	Rainbow trout prep
2	353.5	78.6	115	22 AAU07667	Rainbow trout prep
3	85.5	19.0	114	22 AAU07666	Rainbow trout prep
4	79.5	17.7	116	21 AAG03774	Human secreted pro
5	70.5	15.7	124	22 ABG08924	Novel human diagno
6	70	15.6	117	22 ABG02976	Angiotensin conver
7	69.5	15.4	125	3 AAP20029	Sequence of prepro
8	69	15.3	190	22 AAU55408	Propionibacterium
9	69	15.3	814	22 AAU30069	Novel human secret
10	67.5	15.0	95	22 ABB42770	Peptide #10276 enc

11	67.5	15.0	95	22 ABB26063	Protein #8062 enco
12	67.5	15.0	95	22 AAM36661	Human brain expres
13	67.5	15.0	95	22 AAM76475	Human bone marrow
14	67.5	15.0	95	22 AAM20907	Peptide #7341 enco
15	67.5	15.0	95	22 AAM36583	Peptide #10620 enc
16	67.5	15.0	95	22 ABG45736	Human peptide enco
17	67.5	15.0	179	22 ABG20007	Propionibacterium
18	67	14.9	502	23 ABP51312	Human MDOR SEQ ID
19	67	14.9	1205	16 AAR77363	Endothelial nitrog
20	67	14.9	1205	17 AAR93930	Bovine endothelial
21	66.5	14.8	283	22 AAU01971	Human secreted pro
22	65.5	14.6	146	23 ABG59961	Human DTHP polype
23	64.5	14.3	1305	20 AAW79274	Protein kinase GAK
24	63.5	14.1	333	21 AAY87092	Human secreted pro
25	63.5	14.1	333	22 AAE06069	Human gene 29 enco
26	63.5	14.1	333	23 ABG33891	Human secreted pro
27	63	14.0	56	23 ABP05384	Human OREF protein
28	63	14.0	133	22 ABG27498	Novel human diagno
29	63	14.0	283	23 ABB76961	4-Hydroxyphenylace
30	63	14.0	563	23 ABB76956	Somatostatin-28 (S
31	62.5	13.9	110	23 AAE21871	Propionibacterium
32	62.5	13.9	322	22 AAU44616	Propionibacterium
33	62	13.8	119	21 AAY75023	Neisseria gonorrhe
34	62	13.8	833	20 AAW94058	Murine MSH5
35	61.5	13.7	68	22 AAU40527	Propionibacterium
36	61.5	13.7	92	23 ABP02201	Human OREF protein
37	61.5	13.7	333	22 AAU12237	Human PR04400 poly
38	61.5	13.7	375	21 AAY50971	Human PR04400 prot
39	61.5	13.7	375	21 AAY50971	Neisseria meningit
40	61.5	13.7	376	22 AAM33305	Human polypeptide
41	61.5	13.7	376	23 AAE44445	Human drug metabol
42	61.5	13.7	706	23 ABG27884	Herbicideally activ
43	61	13.6	97	22 AAB50353	Human Ralrinrin pr
44	61	13.6	375	21 AAY75025	Neisseria meningit
45	61	13.6	432	22 ABG02228	Novel human diagno

ALIGNMENTS

RESULT 1

AAU07668

ID AAU07668 standard; Protein; 111 AA.

XX AAU07668;

XX 04-DEC-2001 (first entry)

XX Rainbow trout preprosomatostatin II (PPSS-II'') polypeptide.

XX Rainbow trout: somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
KW neoplasia; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

OS Oncorhynchus mykiss.

XX Key Location/Qualifiers

XX Peptide 1.25

XX /note= "Signal peptide"

XX Protein 1.86

XX /note= "PPSS-II'' pre-sequence"

XX Protein 26.111

XX /note= "Mature PPSS-II''"

XX Peptide 87.97

XX /note= "PPSS-II'' pro-sequence"

XX Peptide 87.111

XX /note= "Prosomatostatin II''"

XX Cleavage-site 96.97

FT Peptide /note- "Dibasic cleavage site"
 FT 98..111
 FT /note- "SS-14 variant peptide"
 PN CA2325169-A1.
 PD 03-JUN-2001.
 PF 01-DEC-2000; 2000CA-2325169.
 PR 03-DEC-1999; 99US-0168934.
 PA (NDSU-) NDSU RES FOUND.
 PI Sheridan MA, Moore CA, Kittelson JD;
 DR WPI: 2001-425997/46.
 DR N-PSDB; AAS12935.
 XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
 FT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
 FT immunodeficiency syndrome and neurological disorders -
 XX
 Claim 1; Fig 3; 52pp; English.
 CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
 CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
 CC preprosomatostatin II (PPSS-II). The protein sequences and their
 CC associated polynucleotides are useful for identifying modified
 CC somatostatin polypeptide which functions as a somatostatin agonist useful
 CC for research, therapeutics or diagnostics, including medical and
 CC veterinary applications. The wild-type somatostatin and its modified
 CC version are useful for treating hypersecretion from endocrine tumours in
 CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
 CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
 CC through their effects on cell proliferation and apoptosis and as adjuncts
 CC in the treatment of diabetes mellitus via inhibition of growth hormone
 CC and glucagon. In addition, dysfunctional somatostatin secretion is
 CC associated with acquired immunodeficiency syndrome (AIDS) and various
 CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
 CC Huntington's disease) and somatostatin antagonists are effective in the
 CC treatment of such conditions. Nucleic acids encoding the polypeptides are
 CC useful in gene therapy and fusion peptides can be targeted to neoplasms
 CC and their metastases, inhibiting the release of their secretory products.
 CC This sequence represents O. mykiss PPSS-II'' protein.
 CC Note: The features for this sequence are specifically claimed in the
 CC specification.
 XX
 SQ Sequence 111 AA:
 Query Match 100.08; Score 450; DB 22; Length 111;
 Best Local Similarity 100.08; Pred. No. 3.2e-49;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MRVSQTHCALALIGLALTCGASQSPDLASRLRLQALAAALPHRSVSRWRTFY 60
 Db. 1 MRVSQTHCALALIGLALTCGASQSPDLASRLRLQALAAALPHRSVSRWRTFY 60
 Qy 61 PNCPLRWPRKVPQPKAKEDLER 86
 Db 61 PNCPLRWPRKVPQPKAKEDLER 86
 RESULT 2
 AAU07667
 ID AAU07667 standard; Protein; 115 AA.
 AC AAU07667;
 XX
 XX AAU07667;
 XX
 XX 04-DEC-2001 (first entry)
 XX
 XX Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
 XX

Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
 PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
 gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
 carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
 glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
 epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
 neoplasm; metastasis; gene therapy; antidiabetic; nontropic; cytostatic;
 anti-human immunodeficiency virus; osteopathic; anticonvulsant.
 Oncorhynchus mykiss:
 Key Location/Qualifiers
 FT Peptide 1..25 "Signal peptide"
 FT Protein 1..87
 FT Protein 26..115 /note- "PPSS-II' pre-sequence"
 FT Misc-difference 74 /note- "Mature PPSS-II'"
 FT /note- "Encoded by CAA"
 FT Peptide 88..101
 FT Peptide 88..115 /note- "PPSS-II' pro-sequence"
 FT /note- "Prosomatostatin II'"
 FT Cleavage-site 100..101
 FT /note- "Dibasic cleavage site"
 FT Peptide 102..115
 FT /note- "SS-14 variant peptide"
 PN CA2325169-A1.
 PD 03-JUN-2001.
 PF 01-DEC-2000; 2000CA-2325169.
 PR 03-DEC-1999; 99US-0168934.
 PA (NDSU-) NDSU RES FOUND.
 PI Sheridan MA, Moore CA, Kittelson JD;
 DR WPI: 2001-425997/46.
 DR N-PSDB; AAS12934.
 XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
 FT for treating diabetes mellitus, acromegaly, gastrinoma, acquired
 FT immunodeficiency syndrome and neurological disorders -
 XX
 Claim 2; Fig 3; 52pp; English.
 CC The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
 CC containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
 CC preprosomatostatin II (PPSS-II). The protein sequences and their
 CC associated polynucleotides are useful for identifying modified
 CC somatostatin polypeptide which functions as a somatostatin agonist useful
 CC for research, therapeutics or diagnostics, including medical and
 CC veterinary applications. The wild-type somatostatin and its modified
 CC version are useful for treating hypersecretion from endocrine tumours in
 CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
 CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
 CC through their effects on cell proliferation and apoptosis and as adjuncts
 CC in the treatment of diabetes mellitus via inhibition of growth hormone
 CC and glucagon. In addition, dysfunctional somatostatin secretion is
 CC associated with acquired immunodeficiency syndrome (AIDS) and various
 CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
 CC Huntington's disease) and somatostatin antagonists are effective in the
 CC treatment of such conditions. Nucleic acids encoding the polypeptides are
 CC useful in gene therapy and fusion peptides can be targeted to neoplasms
 CC and their metastases, inhibiting the release of their secretory products.
 CC This sequence represents O. mykiss PPSS-II'' protein.
 CC Note: The features for this sequence are specifically claimed in the
 CC specification.
 XX
 SQ Sequence 111 AA:
 Query Match 100.08; Score 450; DB 22; Length 111;
 Best Local Similarity 100.08; Pred. No. 3.2e-49;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MRVSQTHCALALIGLALTCGASQSPDLASRLRLQALAAALPHRSVSRWRTFY 60
 Db. 1 MRVSQTHCALALIGLALTCGASQSPDLASRLRLQALAAALPHRSVSRWRTFY 60
 Qy 61 PNCPLRWPRKVPQPKAKEDLER 86
 Db 61 PNCPLRWPRKVPQPKAKEDLER 86
 RESULT 2
 AAU07667
 ID AAU07667 standard; Protein; 115 AA.
 AC AAU07667;
 XX
 XX AAU07667;
 XX
 XX 04-DEC-2001 (first entry)
 XX
 XX Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
 XX

XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasoactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, nonaroidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis.

XX SQ Sequence 517 AA; Query Match 15.6%; Score 70; DB 22; Length 517;
 CC Best Local Similarity 37.8%; Pred. No. 3.7; Mismatches 18; Gaps 4;
 CC Matches 28; Conservative 7;
 QY 14 GLAALICSGAASQPDLDLASRELLORA-----LAAALPHRSGVSEWTFYPCPCCLR 67.
 DB 355 GLAALICSGSPAEK-----KOLLPRAREQEELLAPADGERGSPSFLRHP-----LP 403
 QY 68 W-RPRKVGKPOLKA 80
 DB 404 FPRPRFPSPRLSA 417

RESULT 7
 AAP20029 ID AAP20029 standard; Protein; 125 AA.
 AC AAP20029;
 DT 16-AUG-2002 (updated)
 DT 14-AUG-1992 (first entry)
 XX Sequence of preprosomatostatin-2 encoded on plas2.
 DE Somatostatin; growth hormone; peptide hormone; secretion.
 KW Lophius americanus.
 XX FH Key Location/Qualifiers
 FT Protein 112..125
 FT /label= Somatostatin II
 XX EP46669-A.
 XX 03-MAR-1982.
 XX 21-AUG-1981; 81EP-0303825.
 XX 25-AUG-1980; 80US-0181046.
 XX (REGC) UNIV OF CALIFORNIA.
 XX Hobart P, Crawford R, Pictet RL, Rutter WJ;
 DR WPI: 1982-18113E/10.
 DR N-PSDB; AAN20034.
 XX New somatostatin and precursors - produced by transformed
 PT microorganisms
 XX Example; Fig 3; 50pp; English.
 XX The inventors claim preprosomatostatin-1, prosomatostatin-1,

CC preprosomatostatin-2, prosomatostatin-2 and somatostatin-2; and DNA
 CC encoding them. The translation of somatostatin mRNA yields a
 CC precursor (prepro S1) containing a signal peptide which may be
 CC released during the transit into the endoplasmic reticulum, and the
 CC resultant precursor (pro S1) is subsequently cleaved to yield S1
 CC itself. The prepeptide portion of prepro S1 is probably about 20-25
 CC bases long. Translation of plas2 predicts the sequence of a 125 AA
 CC peptide which surprisingly contains a 14 AA sequence at its carboxy
 CC terminus which differs from S1 by only 2 AAs, and is termed
 CC Somatostatin 2 (S2).
 CC (Updated on 16-AUG-2002 to add missing OS field.)

XX SQ Sequence 125 AA; Query Match 15.4%; Score 69.5; DB 3; Length 125;
 CC Best Local Similarity 36.1%; Pred. No. 0.78; Mismatches 15; Gaps 3;
 CC Matches 22; Conservative 7;
 QY 6 THCALALGLALATCSOGAASQ-----PDLDLASRR--LLORALAAALPHRSGVSR 55
 DB 4 TRCPAILALLALVLCGPSVSSQLDRESDNODLDLELRQHWLLERARSAGL-----LSQE 58
 QY 56 W 56
 DB 59 W 59

RESULT 8
 AAU55408 ID AAU55408 standard; Protein; 190 AA.
 AC AAU55408;
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #16304.
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX Propionibacterium acnes.
 XX WO200181591-A2.
 XX 01-NOV-2001.
 XX 20-APR-2001; 2001WO-US12865.
 XX 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX (CORI-) CORIXA CORP.
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI: 2001-616774/71.
 DR N-PSDB; RAS59569.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

Example 1: SEQ ID No 16603; 1669pp; English..

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 polypeptides. The proteins and their associated DNA sequences are used in
 the treatment, prevention and diagnosis of medical conditions caused by
 P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;
 Best Local Similarity 27.7%; Pred. No. 1;
 Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSOGAASQPDLDLASRRLLQRLAALPHRSVSE 54
 |||||
 Db 2 LALLYLSRVNRLRGPNAYLTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55

QY 55 RW-----RTFYPNCPCLWRPRK 72
 |||||
 Db 56 SWSWDRSWWTRC---AWRPRR 75

RESULT 11

ABB26063
 ID ABB26063 standard; Protein; 95 AA.

XX ABB26063;

XX DT 23-JAN-2002 (first entry)

DE Protein #8062 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234587.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts.

XX Claim 15; SEQ ID NO 27833; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present invention is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;
 Best Local Similarity 27.7%; Pred. No. 1;
 Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSOGAASQPDLDLASRRLLQRLAALPHRSVSE 54
 |||||
 Db 2 LALLYLSRVNRLRGPNAYLTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55

QY 55 RW-----RTFYPNCPCLWRPRK 72
 |||||
 Db 56 SWSWDRSWWTRC---AWRPRR 75

RESULT 12

AM63661
 ID AM63661 standard; Protein; 95 AA.

XX AM63661;

XX DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35766.

XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234587.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.

XX Example 4; SEQ ID NO: 35766; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, f
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX Sequence 95 AA;

Query Match 15.0%; Score 67.5; DB 22; Length 95;
 Best Local Similarity 27.7%; Pred. No. 1;
 Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

QY 10 LALLGLALAI-----CSOGAASQPDLDLASRRLLQRLAALPHRSVSE 54
 |||||
 Db 2 LALLYLSRVNRLRGPNAYLTAGSTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55

Qy 55 RW-----RTFVNCPCLRWRPRK 72
 Db 56 SWSWMDRSWWTTRC---AWRPRR 75

RESULT 13

AAAM76475
 ID AAAM76475 standard; Protein; 95 AA.

XX AC AAAM76475;
 XX XX

DT 06-NOV-2001 (first entry)
 XX XX

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 36781.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

OS WO200157276-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow

XX Example 4; SEQ ID NO: 36781; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia, and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

Sequence 95 AA:

Query Match 15.08; Score 67.5; DB 22; Length 95;

Best Local Similarity 27.7%; Pred. No. 1;

Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

Qy 10 LALLGLALAI-----CSQGAASQPDLDLASRRLQALAAALPHRSGVSE 54

Db 2 LALLYLSLRNLLRGNPAYLNTAGSTLGRSKPSAHFTA-----ASAAASPRSSLRA 55

Qy 55 RW-----RTFVNCPCLRWRPRK 72

Db 56 SWSWMDRSWWTTRC---AWRPRR 75

RESULT 14

AAAM20907
 ID AAAM20907 standard; Protein; 95 AA.

XX AC AAAM20907;
 XX XX

DT

XX 12-OCT-2001 (first entry)

DE Peptide #7341 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer.

OS Homo sapiens.

PN WO200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells

XX Claim 27; SEQ ID NO 25733; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENPs: see AA110068-AA28459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Query Match 15.08; Score 67.5; DB 22; Length 95;

XX Best Local Similarity 27.7%; Pred. No. 1;

XX Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;

Qy 10 LALLGLALAI-----CSQGAASQPDLDLASRRLQALAAALPHRSGVSE 54

Db 2 LALLYLSLRNLLRGNPAYLNTAGSTLGRSKPSAHFTA-----ASAAASPRSSLRA 55

Qy 55 RW-----RTFVNCPCLRWRPRK 72

Db 56 SWSWMDRSWWTTRC---AWRPRR 75

RESULT 15

AAAM36583
 ID AAAM36583 standard; Protein; 95 AA.

XX AC AAAM36583;
 XX XX

DT 17-OCT-2001 (first entry)

DE Peptide #10620 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

```

XX OS Homo sapiens.
XX DN WO200157272-A2
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WI 2001-488897/53.
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta.
XX PS Claim 27; SEQ ID No 36852; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AA131315-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders.
XX SQ Sequence 95 AA;
Query Match 15.0%; Score 67.5; DB 22; Length 95;
Best Local Similarity 27.7%; Pred. No. 1;
Matches 23; Conservative 11; Mismatches 20; Indels 29; Gaps 4;
QY 10 LALLGLALAI-----CSQGAASQPDLDLASRRLLQALAAALPHRSGVSE 54
Db 2 LALLYLSLRVNRLLGNPAYLNTAGTLGRSSKPSAHFTA-----ASAAASPRSSSLRA 55
IY 55 RW-----RTFYPNCPCLRPRK 72
Db 56 SWSWDRSWSWTRC---AWRPRR 75

```

Search completed: March 21, 2003, 11:38:00
Job time : 55.1502 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 2.12446 Seconds
(without alignments)
276.816 Million cell updates/sec

Title: US-09-727-739b-18

Perfect score: 58

Sequence: 1 SVDNPPRRK 11

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries.

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	65.5	33	10 US-09-050-516-46	Sequence 46, Appl
2	38	65.5	197	10 US-09-764-860-343	Sequence 343, App
3	35	60.3	84	10 US-09-864-761-43933	Sequence 43933, A
4	35	60.3	214	12 US-10-062-254-156	Sequence 156, App
5	35	60.3	214	12 US-10-062-254-158	Sequence 158, App
6	35	60.3	398	9 US-09-291-417-30	Sequence 30, Appl
7	35	60.3	409	9 US-09-533-029-104	Sequence 104, App
8	35	60.3	591	9 US-09-291-417-103	Sequence 103, App
9	35	60.3	751	10 US-09-815-242-14001	Sequence 14001, A
10	34	58.6	215	10 US-09-050-516-47	Sequence 47, Appl
11	34	58.6	293	10 US-09-815-242-5614	Sequence 5614, Ap
12	34	58.6	299	10 US-09-815-242-10205	Sequence 10205, A
13	34	58.6	303	9 US-09-828-523A-20	Sequence 20, Appl
14	34	58.6	303	9 US-09-815-242-12527	Sequence 12527, A
15	34	58.6	311	9 US-09-828-523A-86	Sequence 86, Appl
16	34	58.6	528	10 US-09-792-2008-22	Sequence 22, Appl
17	34	58.6	679	10 US-09-050-516-42	Sequence 42, Appl
18	33	56.9	160	10 US-09-916-790-35	Sequence 35, Appl
19	33	56.9	206	9 US-09-738-626-6537	Sequence 6537, Ap

20	33	56.9	265	10 US-09-925-122A-1	Sequence 1, Appli
21	33	56.9	280	12 US-10-115-899-5	Sequence 5, Appli
22	33	56.9	280	12 US-10-115-899-8	Sequence 8, Appli
23	33	56.9	285	10 US-09-815-242-4869	Sequence 4869, Ap
24	33	56.9	334	10 US-09-815-242-10798	Sequence 10798, A
25	33	56.9	345	9 US-10-036-729-4	Sequence 4, Appli
26	33	56.9	547	10 US-09-815-242-11982	Sequence 11982, A
27	33	56.9	683	9 US-10-025-380-122	Sequence 122, App
28	33	56.9	683	10 US-09-922-217-122	Sequence 122, App
29	33	56.9	683	10 US-09-833-263-122	Sequence 122, App
30	33	56.9	764	10 US-09-866-583-41	Sequence 41, Appl
31	32	55.2	77	9 US-10-091-483-174	Sequence 174, Appl
32	32	55.2	77	10 US-09-764-846-174	Sequence 174, App
33	32	55.2	84	9 US-10-091-483-247	Sequence 247, App
34	32	55.2	84	10 US-09-764-846-247	Sequence 247, App
35	32	55.2	251	10 US-09-768-826-35	Sequence 35, Appl
36	32	55.2	296	10 US-09-815-242-13561	Sequence 13561, A
37	32	55.2	314	10 US-09-768-826-54	Sequence 54, Appl
38	32	55.2	315	9 US-10-012-342-297	Sequence 297, App
39	32	55.2	431	9 US-10-012-342-146	Sequence 146, App
40	32	55.2	678	9 US-09-738-626-6025	Sequence 6025, Ap
41	32	55.2	713	10 US-09-802-741A-3	Sequence 3, Appli
42	32	55.2	715	10 US-09-431-226-2	Sequence 2, Appli
43	32	55.2	1596	9 US-09-909-567B-47	Sequence 47, Appli
44	31.5	54.3	842	10 US-09-798-831-8	Sequence 8, Appli
45	31	53.4	86	10 US-09-864-761-46557	Sequence 46557, A

ALIGNMENTS

RESULT 1

US-09-050-516-46
Sequence 46, Application US/09050516

Patent No. US20010010904A1

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: HAYDEN, MARK

APPLICANT: HODGES, STEVEN C.

APPLICANT: KASS, MICHAEL R.

APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

FOR DETECTING DISEASES OF THE GASTROINTESTINAL

TITLE OF INVENTION: TRACT

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/050.516

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/828,855

FILING DATE: 31-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.
 REGISTRATION NUMBER: 35,441
 REFERENCE/DOCKET NUMBER: 6065, US, P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847/935-1729
 TELEFAX: 847/938-2623
 TELEX:
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 33 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. US20010010904A1e
 US-09-050-516-46

Query Match 65.5%; Score 38; DB 10; Length 33;
 Best Local Similarity 63.6%; Pred. No. 2.3;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEK 11
 ||||| |||||
 nb 3 SVDSAPPGQK 13

RESULT 2

US-09-764-860-343
 Sequence 343, Application US/09764860
 Patent No. US20020094953A1
 GENERAL INFORMATION:

APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PC008
 CURRENT APPLICATION NUMBER: US/09/764,860
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PALM or file wrapper
 NUMBER OF SEQ ID NOS: 1198
 SOFTWARE: PatentIn ver. 2.0
 SEQ ID NO 343
 LENGTH: 197

TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (53)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (97)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-764-860-343

Query Match 65.5%; Score 38; DB 10; Length 197;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDNLP 7
 |||||
 Db 174 SVDNLP 180

RESULT 3

US-09-864-761-43933
 Sequence 43933, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aesomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 43933
 LENGTH: 84
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC005392.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
 OTHER INFORMATION: SWISSPROT HIT: Q00888, EVALUATE 3.00e-46
 US-09-864-761-43933

Query Match 60.3%; Score 35; DB 10; Length 84;
 Best Local Similarity 54.5%; Pred. No. 23;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPEK 11
 :|||: |||||
 Db 8 TINNLNPREK 18

RESULT 4

US-10-062-254-156
 Sequence 156, Application US/10062254
 Patent No. US20020138882A1
 GENERAL INFORMATION:

APPLICANT: Cahoon, Edgar B
 APPLICANT: Cahoon, Rebecca E
 APPLICANT: Falco, Saverio Carl
 APPLICANT: Fang, Yiven
 APPLICANT: Hantke, Sabine S.
 APPLICANT: Lee, Jian-Ming

APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE: US/10/062,254

CURRENT APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 2002-02-01

PRIOR FILING DATE: 2000-07-28

PRIOR FILING DATE: 2000-07-28

PRIOR FILING DATE: 1999-07-30

PRIOR FILING DATE: 1999-09-23

PRIOR FILING DATE: 1999-09-30

PRIOR FILING DATE: 1999-10-01

PRIOR FILING DATE: 1999-12-09

PRIOR FILING DATE: 1999-12-16

PRIOR FILING DATE: 1999-12-21

PRIOR FILING DATE: 1999-12-22

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 375

SOFTWARE: Microsoft Office 97

SEQ ID NO 156

LENGTH: 214

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: UNSURE

LOCATION: (107)

US-10-062-254-156

Query Match 60.3%; Score 35; DB 12; Length 214;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPFR 10

DB 9 SLDHLPSEQ 18

RESULT 5

US-10-062-254-158

Sequence 158, Application US/10062254

Patent No. US2002013882A1

GENERAL INFORMATION:

APPLICANT: Caboon, Edgar B

APPLICANT: Caboon, Rebecca E

APPLICANT: Falco, Savario Carl

APPLICANT: Fang, Yixue

APPLICANT: Hantke, Sabine S

APPLICANT: Lee, Jian-Ming

APPLICANT: Li, Zhongsen

APPLICANT: Miao, Guo-Hua

APPLICANT: Morgante, Michele

APPLICANT: Niu, Xiping

APPLICANT: Odell, Joan

APPLICANT: Rafalski, Antoni

APPLICANT: Sakai, Hajime

APPLICANT: Zheng, Peizhong

APPLICANT: Zhu, Qun

TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism

FILE REFERENCE: US/10/062,254

CURRENT APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 2002-02-01

PRIOR FILING DATE: 2000-07-28

PRIOR FILING DATE: 2000-07-28

PRIOR FILING DATE: 1999-07-30

PRIOR FILING DATE: 1999-09-23

PRIOR FILING DATE: 1999-09-30

PRIOR FILING DATE: 1999-10-01

PRIOR FILING DATE: 1999-12-09

PRIOR FILING DATE: 1999-12-16

PRIOR FILING DATE: 1999-12-21

PRIOR FILING DATE: 1999-12-22

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 375

SOFTWARE: Microsoft Office 97

SEQ ID NO 158

LENGTH: 214

TYPE: PRT

ORGANISM: Glycine max

US-10-062-254-158

Query Match 60.3%; Score 35; DB 12; Length 214;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPFR 10

DB 9 SLDHLPSEQ 18

RESULT 6

US-09-291-417-30

Sequence 30, Application US/09291417A

Publication No. US20030050230A1

GENERAL INFORMATION:

APPLICANT: PLOWMAN, GREGORY

APPLICANT: MARTINEZ, RICARDO

APPLICANT: WHYTE, DAVID

TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

FILE REFERENCE: 240/300

CURRENT APPLICATION NUMBER: US/09/291,417A

CURRENT FILING DATE: 1999-04-13

EARLIER FILING DATE: 1999-04-13

EARLIER FILING DATE: 1998-04-14

NUMBER OF SEQ ID NOS: 147

SOFTWARE: PastSeq for Windows Version 3.0

SEQ ID NO 30

LENGTH: 398

TYPE: PRT

ORGANISM: Mammalian (Human) PAK5

US-09-291-417-30

Query Match 60.3%; Score 35; DB 9; Length 398;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPR 8

DB 336 DNLPPR 341

RESULT 7

US-09-533-029-104
; Sequence 104, Application US/09533029
; Publication No. US20030046723A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Broun, Pierre
; APPLICANT: Kiedde, James
; APPLICANT: Pineda, Omaia
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1034
US-09-533-029-104

Query Match 60.3%; Score 35; DB 9; Length 409;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0;

QY 4 NLPKR 10
Db 86 NLPKR 92

RESULT 8
US-09-291-417-103
; Sequence 103, Application US/09291417A
; Publication No. US20030050230A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 240/300
; CURRENT APPLICATION NUMBER: US/09/291,417A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: US 60/081,784
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 103
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Full Length Mammalian (Human) PAK5hu
US-09-291-417-103

Query Match 60.3%; Score 35; DB 9; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 3 DNLPR 8
Db 529 DNLPR 534

RESULT 9
US-09-815-242-14001
; Sequence 14001, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14001
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14001

Query Match 60.3%; Score 35; DB 10; Length 751;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPR 9
Db 742 DNLPR 748

RESULT 10
US-09-050-516-47
; Sequence 47, Application US/09050516
; Patent No. US20010010904A1
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HAYDEN, MARK
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE GASTROINTESTINAL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,516
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/828,855
FILING DATE: 31-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6065 US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. US20010010904A1e
US-09-050-516-47

Query Match 58.6%; Score 34; DB 10; Length 215;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VDNLPPEK 11
||: ||: ||
Db 160 VDSAPPQGRK 169

RESULT 11

US-09-815-242-5614
Sequence 5614, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5614
LENGTH: 293
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5614

Query Match 58.6%; Score 34; DB 10; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VDNLPP 7
|||||
Db 33 VDNLPP 38

RESULT 12

US-09-815-242-10205
Sequence 10205, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 10205
LENGTH: 299
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10205

Query Match 58.6%; Score 34; DB 10; Length 299;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SVDNLPPE 9
|||||
Db 15 STDNLPLE 23

RESULT 13

US-09-828-523A-20
Sequence 20, Application US/09828523A
Patent No. US20020168697A1
GENERAL INFORMATION:

APPLICANT: The Pharmacia & Upjohn Company.
TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
FILE REFERENCE: 268, 62120101
CURRENT APPLICATION NUMBER: US/09/828, 523A
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/266, 327
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 99
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 303
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-828-523A-20

Query Match. 58.6%; Score 34; DB 9; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VDNLPP 7
| | | | |
Db 39 VDNLPP 44

SULT 14
US-09-815-242-12527
Sequence 12527, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
PROKARYOTES
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27/257,931
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12527
LENGTH: 303
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12527

Query Match. 58.6%; Score 34; DB 10; Length 303;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VDNLPP 7
| | | | |
Db 39 VDNLPP 44

RESULT 15
US-09-828-523A-86
Sequence 86, Application US/09828523A
Patent No. US20020168697A1
GENERAL INFORMATION:
APPLICANT: The Pharmacia & Upjohn Company
TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
FILE REFERENCE: 268, 62120101
CURRENT APPLICATION NUMBER: US/09/828, 523A
CURRENT FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/266, 327
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 99
SOFTWARE: PatentIn version 3.1
SEQ ID NO 86
LENGTH: 311
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned
in E. coli.
US-09-828-523A-86

Query Match. 58.6%; Score 34; DB 9; Length 311;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VDNLPP 7
| | | | |
Db 39 VDNLPP 44

Search completed: March 21, 2003, 11:56:31
Job time: 4.12446 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: March 21, 2003, 11:41:02 ; Search time 2.07725 Seconds
(without alignments)
155.808 Million cell updates/sec

Title: US-09-727-739B-18
Perfect score: 58
Sequence: 1 SVDNLPPEPRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCUTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	67.2	219	4	US-09-247-737B-52
2	39	67.2	224	4	US-09-247-737B-34
3	39	67.2	248	4	US-09-134-001C-5085
4	37	63.8	424	6	5169835-6
5	36	62.1	273	4	US-09-134-001C-3917
6	36	62.1	2496	4	US-09-125-028-2
7	36	62.1	2958	4	US-08-894-344C-2
8	35	60.3	230	6	5169835-13
9	35	60.3	419	6	5169835-2
10	35	60.3	521	3	US-09-082-737-2
11	34	58.6	223	3	US-08-857-534-12
12	34	58.6	223	5	PCT-US95-04971-12
13	34	58.6	229	4	US-09-247-737B-48
14	34	58.6	305	4	US-09-134-001C-5632
15	34	58.6	361	1	US-08-258-261B-3
16	34	58.6	361	1	US-08-456-837-3
17	34	58.6	361	1	US-08-457-342-3
18	34	58.6	361	1	US-08-457-646A-3
19	34	58.6	361	1	US-08-457-076A-3
20	34	58.6	361	1	US-08-457-335A-3
21	34	58.6	361	2	US-08-729-214-3
22	34	58.6	361	3	US-09-028-934-3
23	33	56.9	180	4	US-09-134-001C-5294
24	33	56.9	265	2	US-08-970-133-1
25	33	56.9	265	4	US-09-294-545-1
26	33	56.9	280	4	US-09-601-478-5
27	33	56.9	280	4	US-09-601-478-8

28	33	56.9	345	1	US-08-031-148-4	Sequence 4, Appli
29	33	56.9	345	3	US-08-415-838-4	Sequence 4, Appli
30	33	56.9	345	4	US-09-205-169-4	Sequence 4, Appli
31	33	56.9	415	2	US-08-576-626A-52	Sequence 52, Appli
32	33	56.9	683	1	US-07-878-960-2	Sequence 2, Appli
33	33	56.9	855	1	US-08-344-536-2	Sequence 2, Appli
34	33	56.9	855	1	US-08-344-536-5	Sequence 5, Appli
35	33	56.9	855	3	US-08-920-562-2	Sequence 2, Appli
36	33	56.9	855	3	US-08-920-562-5	Sequence 5, Appli
37	32	55.2	2544	2	US-08-576-626A-32	Sequence 32, Appli
38	32	55.2	275	4	US-09-147-915-4	Sequence 4, Appli
39	32	55.2	296	3	US-08-986-769-2	Sequence 2, Appli
40	32	55.2	361	3	US-09-028-934-30	Sequence 30, Appli
41	32	55.2	449	1	US-08-624-663A-2	Sequence 2, Appli
42	32	55.2	449	2	US-08-974-565C-1	Sequence 1, Appli
43	32	55.2	449	3	US-09-255-748-1	Sequence 1, Appli
44	32	55.2	470	4	US-09-004-838-44	Sequence 44, Appli
45	32	55.2	472	4	US-09-004-838-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1
US-09-247-737B-52
; Sequence 52, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; CURRENT FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 52
; LENGTH: 219
; TYPE: PRT
; ORGANISM: SOYBEAN
US-09-247-737B-52

Query Match 67.2%; Score 39; DB 4; Length 219;
Best Local Similarity 75.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
Db 194 DNLPPRRK 201

RESULT 2
US-09-247-737B-34
; Sequence 34 Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1108-A
; CURRENT APPLICATION NUMBER: US/09/247,373B
; CURRENT FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 08/924,747
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 34
; LENGTH: 224
; TYPE: PRT
; ORGANISM: SOYBEAN
US-09-247-737B-34

```
Query Match          57.2%; Score 39; DB 4; Length 224;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
Db 202 DNLPPRDK 209

RESULT 3
US-09-134-001C-5085
; Sequence 5085, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5085
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5085

Query Match          67.2%; Score 39; DB 4; Length 248;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVDNLPPIRR 10
Db 181 SAKNLPPIRR 190

RESULT 4
5169835-6
; Patent No. 5169835
; APPLICANT: WAI-YEE, CHAN
; TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,409
; FILING DATE: 07-AUG-1989
; SEQ ID NO: 6
; LENGTH: 424
5169835-6

Query Match          63.8%; Score 37; DB 6; Length 424;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVDNLPPIRR 11
Db 244 TINNLPPIRR 254

RESULT 5
US-09-134-001C-3917
; Sequence 3917, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
```

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3917
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3917
```

```
Query Match          62.1%; Score 36; DB 4; Length 273;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 VDNLPPIRR 11
Db 40 VKHLPPIRR 49
```

```
RESULT 6
US-09-125-028-2
; Sequence 2, Application US/09125028A
; Patent No. 6190707
; GENERAL INFORMATION:
; APPLICANT: WADDOUX, Isabelle
; APPLICANT: COLAVIZZA, Didier
; APPLICANT: LOIEZ, Annie
; TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS
; FILE REFERENCE: leuvre sensible froid
; CURRENT APPLICATION NUMBER: US/09/125,028A
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: PCT/FR97/00254
; EARLIER FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2496
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-125-028-2
```

```
Query Match          62.1%; Score 36; DB 4; Length 2496;
Best Local Similarity 66.7%; Pred. No. 4,4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 VDNLPPIRR 10
Db 257 IDVLPPIRR 265
```

```
RESULT 7
US-08-894-344C-2
; Sequence 2, Application US/08894344C
; Patent No. 6172196
; GENERAL INFORMATION:
; APPLICANT: KAWASAKI, Hideki
; APPLICANT: TOKAI, Masaya
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: OUCHI, KOZO
; TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING
; TITLE OF INVENTION: YEAST
; NUMBER OF SEQUENCES: 2
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Diskette - 3.50 inch, 1440 Kb storage.

COMPUTER: IBM PS/V

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:13

LENGTH: 230

5169835-13

Query Match

Best Local Similarity

Score 35; DB 6; Length 230;

Pred. No. 53;

Mismatches 3; Indels 2; Gaps 0;

Matches 6; Conservative 3;

QY 1 SVDNLPKPK 11

DB 101 TINLAPRENK 111

RESULT 9

5169835-2

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:2

LENGTH: 419

5169835-2

Query Match

Best Local Similarity

Score 35; DB 6; Length 230;

Pred. No. 53;

Mismatches 3; Indels 2; Gaps 0;

Matches 6; Conservative 3;

QY 1 SVDNLPKPK 11

DB 101 TINLAPRENK 111

RESULT 9

5169835-2

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:2

LENGTH: 419

5169835-2

Query Match

Best Local Similarity

Score 35; DB 6; Length 230;

Pred. No. 53;

Mismatches 3; Indels 2; Gaps 0;

Matches 6; Conservative 3;

QY 1 SVDNLPKPK 11

DB 101 TINLAPRENK 111

RESULT 9

5169835-2

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:2

LENGTH: 419

5169835-2

Query Match

Best Local Similarity

Score 35; DB 6; Length 230;

Pred. No. 53;

Mismatches 3; Indels 2; Gaps 0;

Matches 6; Conservative 3;

QY 1 SVDNLPKPK 11

DB 101 TINLAPRENK 111

RESULT 9

5169835-2

APPLICANT: WAI-YEE, CHAN

TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS

NUMBER OF SEQUENCES: 48

CURRENT APPLICATION NUMBER: US/07/390,409

FILING DATE: 07-AUG-1989

SEQ ID NO:2

LENGTH: 419

Query Match

Best Local Similarity

Score 35; DB 6; Length 419;

Pred. No. 1e+02;

Mismatches 3; Indels 2; Gaps 0;

Matches 6; Conservative 3;

QY 1 SVDNLPKPK 11

DB 245 TINLAPRENK 255

RESULT 10

US-09-082-737-2

Sequence 2, Application US/09082737

Patent No. 6013500

GENERAL INFORMATION:

APPLICANT: Minden, Audrey

TITLE OF INVENTION: PAK4; A NO. 6013500e1 Gene Encoding A Serine/

TITLE OF INVENTION: Threonine Kinase

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 11230

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentcin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/082,737

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/55311

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 591 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-082-737-2

Query Match

Best Local Similarity

Score 35; DB 3; Length 591;

Pred. No. 1.4e+02;

Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0;

QY 3 DNLPPR 8

DB 529 DNLPPR 534

RESULT 11

US-08-857-534-12

Sequence 12, Application US/08857534

Patent No. 6087170

GENERAL INFORMATION:

APPLICANT: George W. Kemble

TITLE OF INVENTION: A No. 6087170e1 VZV Gene, Mutant VZV and Immunogenic

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooley Godward Castro Huddleson & Tatum

STREET: 5 Palo Alto Square

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentcin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/082,737

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/55311

ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/857,534
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/235,406
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cserr
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-004/000S
TELEPHONE: 415-843-5165
TELEFAX: 415-857-0663
TELEX: 380816 CooleyPA
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
MOLECULE TYPE: Protein
US-08-857-534-12

Query Match 58.6%; Score 34; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPRK 11
DB 44 PPRK 49

RESULT 12
PCT-US95-04971-12
US-09-727-739b-18.ra
SEQUENCE 12, Application PC/TUS9504971
GENERAL INFORMATION:
APPLICANT: George W. Kemble
TITLE OF INVENTION: A Novel VZV Gene, Mutant VZV and Immunogenic
COMPOSITIONS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04971
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,406
FILING DATE: APRIL 28, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cserr
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: AVIR-004/000W
TELEPHONE: 415-843-5165
TELEFAX: 415-857-0663
TELEX: 380816 CooleyPA

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: Protein
PCT-US95-04971-12

Query Match 58.6%; Score 34; DB 5; Length 223;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPRK 11
DB 44 PPRK 49

RESULT 13
US-09-247-373B-48
SEQUENCE 48, Application US/09247373B
PATENT NO. 6168954
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 48
LENGTH: 229
TYPE: PRT
ORGANISM: SOYBEAN
US-09-247-373B-48

Query Match 58.6%; Score 34; DB 4; Length 229;
Best Local Similarity 62.5%; Pred. No. 80;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DNLPPRR 10
DB 200 DSLPPRRK 207

RESULT 14
US-09-134-001C-5632
SEQUENCE 5632, Application US/09134001C
PATENT NO. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5632
LENGTH: 305
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5632

Query Match 58.6%; Score 34; DB 4; Length 305;
Best Local Similarity 100.0%; Pred. No. 11e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDLPP 7
Db 42 VDLPP 47

RESULT 15
US-08-258-261B-3
; Sequence 3, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas M.
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Ukes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-258-261B-3

Query Match 58.6%; Score 34; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 NLPPE 9
Db 107 NLPPE 112

Search completed: March 21, 2003, 11:42:35
Job time : 4.07725 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      March 21, 2003, 11:38:15 ; Search time 12.2318 Seconds
           (without alignments)
           421.131 Million cell updates/sec
```

Title: US-09-727-739B-19

Perfect score: 122

Sequence: 1 MRVSIHCALALLGLAICSQGA 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

- Maximum Match 100%
- Listing first 45 summaries

Database :

SPTRMBL_21:*

1:	sp:archea:*
2:	sp:bacteria:*
3:	sp:fungi:*
4:	sp:human:*
5:	sp:invertebrate:*
6:	sp:mammal:*
7:	sp:mhc:*
8:	sp:organelle:*
9:	sp:phase:*
10:	sp:plant:*
11:	sp:rodent:*
12:	sp:virus:*
13:	sp:vertebrate:*
14:	sp:unclassified:*
15:	sp:rvirus:*
16:	sp:bacteriap:*
17:	sp:archeoap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	82	67.2	120	13	Q90Y39	Q90Y39	catostomus
2	76	62.3	114	13	Q90Y42	Q90Y42	patodon bu
3	67	54.9	114	13	Q90Y41	Q90Y41	gnathomemus
4	67	54.9	115	13	Q90Y43	Q90Y43	osteoGLOSSu
5	61	50.0	116	13	Q90XEL	Q90XEL	acipenser t
6	56	45.9	114	13	Q90Y40	Q90Y40	chittala chi
7	54	44.3	216	5	Q96316	Q96316	aphanarthr
8	53	43.4	434	5	Q96656	Q96656	penaeus mon
9	52	42.6	132	5	Q96312	Q96312	hyलगomotu
10	52	42.6	133	5	Q963H3	Q963H3	scolytus un
11	52	42.6	146	5	Q96313	Q96313	xylechinoso
12	52	42.6	146	5	Q96126	Q96126	dendroctonu
13	52	42.6	148	5	Q96310	Q96310	pseudohyles
14	52	42.6	148	5	Q962A0	Q962A0	hylastes po
15	52	42.6	149	5	Q96311	Q96311	dendroctonu
16	52	42.6	149	5	Q96124	Q96124	dendroctonu
17	52	42.6	149	5	Q96124	Q96124	dendroctonu

ALIGNMENTS

RESULT 1

Q90Y39	PRELIMINARY;	PRT;	120 AA.
ID	Q90Y39		
AC	Q90Y39		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Preprosomatostatin.		
DE	Peptidomastatin.		
DE	Catostomus commersoni (White sucker).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinodontiformes; Catostomidae; Catostomus.		
OC	NCBI_TaxId=7971;		
OC	[1]		
RP	SEQUENCE FROM N.A.		
RP	Al-Mahrouki A.A., Irwin D.M., Youson J.H.;		
RA	"Molecular cloning and characterization of white sucker preprosomatostatin".		
RT	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		
RT	EMBL: AF293654; AAK97071.1; .		
DR	InterPro: IPR004250; Somatostatin.		
DR	Pfam: PF03002; Somatostatin.		
DR	SEQUENCE 120 AA: 13783 MW: 0042803526358805		
DR	CRC64:		

RESULT 2

RESOL 2
Q90Y42
ID Q90Y42 PRELIMINARY; PRT; 114 AA.
AC Q90Y42;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Preprosomatostatin.
 OS Pantodon buchholzi (Butterflyfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
 OC Osteoglossiformes; Pantodontidae; Pantodon.
 OX NCBI_TaxID=8276;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
 RT "Characterization of variant somatostatin cDNAs from several
 osteoglossomorphs: molecular identification and comparative
 analysis.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF292651; AAK97068.1;
 DR InterPro: IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 'SQ SEQUENCE 114 AA; 12352 MW; 7E3D44CB6A27B12F CRC64;

Query Match 62.3%; Score 76; DB 13; Length 114;
 Best Local Similarity 52.0%; Pred. No. 0.00095;
 Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

1 MRYVSIHICALALLGLALALICSGAA 25
 | : : | | | | | | | | | | : : |
 Db 1 MKLCVHCILALLGLVLCGSSSA 25

RESULT 3

Q90Y41 ID Q90Y41 PRELIMINARY; PRT; 114 AA.
 AC Q90Y41;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Preprosomatostatin.
 OS Gnathonemus petesili.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
 OC Osteoglossiformes; Mormyridae; Gnathonemus.
 OX NCBI_TaxID=42645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
 RT "Characterization of variant somatostatin cDNAs from several
 osteoglossomorphs: molecular identification and comparative
 analysis.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF292652; AAK97069.1;
 DR InterPro: IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 'SQ SEQUENCE 114 AA; 12494 MW; 454DA57A309CA8F2 CRC64;

Query Match 54.9%; Score 67; DB 13; Length 114;
 Best Local Similarity 64.0%; Pred. No. 0.02;
 Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRYVSIHICALALLGLALALICSGAA 25
 | : : | | | | | | | | | | : : |
 Db 1 MLSSRIQCALALLSLALAVSSVSA 25

RESULT 4

Q90Y43 ID Q90Y43 PRELIMINARY; PRT; 115 AA.
 AC Q90Y43;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Preprosomatostatin.
 OS Osteoglossum bicirrhosum (silver arowana).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;

OC Osteoglossiformes; Osteoglossidae; Osteoglossum.
 OX NCBI_TaxID=109271;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Al-Mahrouki A.A., Irwin D.M., Youson J.H.;
 RT "Characterization of variant somatostatin cDNAs from several
 osteoglossomorphs: molecular identification and comparative
 analysis.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF292650; AAK97067.1;
 DR InterPro: IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 'SQ SEQUENCE 115 AA; 12791 MW; D65FBD7C6F1E4E4D CRC64;

Query Match 54.9%; Score 67; DB 13; Length 115;
 Best Local Similarity 52.0%; Pred. No. 0.02;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 MRYVSIHICALALLGLALALICSGAA 25
 | : : | | | | | | | | | | : : |
 Db 1 MKICQIHCTVLVLLGLVLCPSAA 25

RESULT 5

Q90XEL ID Q90XEL PRELIMINARY; PRT; 116 AA.
 AC Q90XEL;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Somatostatin.
 OS Acipenser transmontanus (White sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=7904;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tabbucchi M., Tostivint H., Lohmann I., Sollars C., Vallarino M.,
 RA Dore R.M., Vaudry H.;
 RT "Polygenic expression of somatostatin in the sturgeon Acipenser
 transmontanus: molecular cloning and distribution of the mRNAs
 encoding two somatostatin precursors.";
 RL J. Comp. Neurol. 0:0-0(2001).
 DR EMBL; AF395849; AAL13248.1;
 DR InterPro: IPR004250; Somatostatin.
 DR Pfam; PF03002; Somatostatin; 1.
 'SQ SEQUENCE 116 AA; 12616 MW; 72E0C3FF6C80650F CRC64;

Query Match 50.0%; Score 61; DB 13; Length 116;
 Best Local Similarity 56.0%; Pred. No. 0.15;
 Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 MRYVSIHICALALLGLALALICSGAA 25
 | : : | | | | | | | | | | : : |
 Db 1 MLSSRIQCALALLSLALAVSSVSA 25

RESULT 6

Q90Y40 ID Q90Y40 PRELIMINARY; PRT; 114 AA.
 AC Q90Y40;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Preprosomatostatin.
 OS Chitala chitala (clown knife-fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
 OC Osteoglossiformes; Notopteridae; Chitala.
 OX NCBI_TaxID=112163;
 RN [1]

```

RP SEQUENCE FROM N.A.
RA Al-Mehrouki A.A., Irwin D.M., Youson J.H.;
RT "Characterization of variant somatostatin cDNAs from several
RT osteoglossomorphs: molecular identification and comparative
RT analysis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF292853; AK97070.1; -
DR InterPro: IPR0004250; Somatostatin; 1.
DR Pfam: PF03002; Somatostatin; 1.
SQ SEQUENCE 114 AA; 12561 MW; 4E3C32F58E3F971 CRC64;

Query Match 45.9%; Score 56; DB 13; Length 114;
Best Local Similarity 56.0%; Pred. No. 0.82;
Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MRVSQIHCALALGLALALICSGAA 25
I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Jb 1 MLSTRIQCALALSLALPVSYYAA 25

RESULT 7
ID Q96316 PRELIMINARY; PRT; 216 AA.
AC Q96316;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENO1N1.
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Aphanarthrum.
OX NCBI_TaxID=113463;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
RT and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375332; AAK54790.1; -
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1 216
FT NON_TER 216 216
SQ SEQUENCE 216 AA; 22700 MW; 34E3408775FED864 CRC64;

Query Match 44.3%; Score 54; DB 5; Length 216;
Best Local Similarity 58.8%; Pred. No. 2.8;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ALALLGLALALICSGAA 25
I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 79 ATAILGVSLAVCKAGAA 95

RESULT 8
ID Q96656 PRELIMINARY; PRT; 434 AA.
AC Q96656;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
OS Penaeus monodon (Penaeid shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OX NCBI_TaxID=6687;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-ABDOMINAL MUSCLES;
RA Boonchuay C., Boonyawan B., Panyim S., Sonthayanon B.;
RT "A cDNA sequence of phosphopyruvate hydratase (enolase) from Black
RT Tiger Prawn, Penaeus monodon.";
RL Asia Pac. J. Mol. Biol. Biotechnol. 7:89-94(1999).
CC -1 CATALYTIC ACTIVITY: 2-PHOSPHO-D-GLYCERATE -> PHOSPHOENOLPYRUVATE +
CC H(2O).
CC -1 COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER (BY SIMILARITY).
CC -1 PATHWAY: GLYCOLYSIS.
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
DR EMBL: AF100985; AAC78141.1; -
DR HSSP: P56252; LPDZ.
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR PRINTS: PR00148; ENOLASE.
DR ProDom: PD000902; Enolase; 1.
DR TIGRFAMs: TIGR01060; eno; 1.
DR PROSITE: PS00164; ENOLASE; 1.
KW Glycolysis; Lyase; Magnesium; Pyruvate.
SQ SEQUENCE 434 AA; 47265 MW; EB575C6F541ABD3 CRC64;

Query Match 43.4%; Score 53; DB 5; Length 434;
Best Local Similarity 64.7%; Pred. No. 7.4;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ALALLGLALALICSGAA 25
I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 ANAILGVSLAICKAGAA 124

RESULT 9
ID Q96312 PRELIMINARY; PRT; 132 AA.
AC Q96312;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENO1N1.
OS Hylurgonotus tuberculatus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Hylurgonotus.
OX NCBI_TaxID=141164;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RA Jordal B.H.;
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
RT and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375337; AAK54795.1; -
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1 132
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 13880 MW; BCB814053FFA3471 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 132;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ALALLGLALALICSGAA 25
I : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 ANAILGVSLAVCKAGAA 36

RESULT 10

```

```

Q963H3          PRELIMINARY;          PRT; 133 AA.
ID Q963H3;
AC Q963H3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Cucujiformia; Phytophaga; Scolytidae; Scolytinae.
OC NCBI_TaxID=159924;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375346; AAK54804.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 133 AA; 13879 MW; BF8D60C5D97AC008 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 133;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 1 ANAILGVSLAVCKAGAA 17

RESULT 11
Q963I3          PRELIMINARY;          PRT; 146 AA.
ID Q963I3;
AC Q963I3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Xylechinomus validianus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Xylechinomus.
OC NCBI_TaxID=141197;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375336; AAK54794.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR PRINTS: PR01610; CD36ANTIGEN.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 146 AA; 15404 MW; E3346CF20ADAF658 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 1 ANAILGVSLAVCKAGAA 17

Q963H3          PRELIMINARY;          PRT; 133 AA.
ID Q963H3;
AC Q963H3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Cucujiformia; Phytophaga; Scolytidae; Scolytinae.
OC NCBI_TaxID=159924;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375346; AAK54804.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 133 AA; 13879 MW; BF8D60C5D97AC008 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 133;
Best Local Similarity 58.8%; Pred. No. 3.6;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 1 ANAILGVSLAVCKAGAA 17

Q963I3          PRELIMINARY;          PRT; 146 AA.
ID Q963I3;
AC Q963I3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Xylechinomus validianus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Xylechinomus.
OC NCBI_TaxID=141197;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375336; AAK54794.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR PRINTS: PR01610; CD36ANTIGEN.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 146 AA; 15404 MW; E3346CF20ADAF658 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 1 ANAILGVSLAVCKAGAA 17

Q963H3          PRELIMINARY;          PRT; 146 AA.
ID Q963H3;
AC Q963H3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Dendroctonus mexicanus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Dendroctonus.
OC NCBI_TaxID=77163;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D.,
RT "Evolutionary origins of Gondwanan interactions: How old are Araucaria
beetle herbivores?";
RL Biol. J. Linn. Soc. 0:0-0(2001).
DR EMBL: AY040304; AAK77977.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 146 AA; 15194 MW; 28D98F57787CC168 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 146;
Best Local Similarity 58.8%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 13 ANAILGVSLAVCKAGAA 29

Q963I3          PRELIMINARY;          PRT; 148 AA.
ID Q963I3;
AC Q963I3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enolase (Fragment).
GN ENOINI.
OS Pseudohylesinus nebulosus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Scolytidae; Pseudohylesinus.
OC NCBI_TaxID=77155;
[1]
RN SEQUENCE FROM N.A.
RA Farrell B.D., Sequeira A.S., O'Meara B.C., Normark B.B., Chung J.H.,
RT "The evolution of agriculture in beetles (Curculionidae: Scolytinae
and Platypodinae).";
RL Evolution 0:0-0(2001).
DR EMBL: AF375339; AAK54797.1;
DR InterPro: IPR000941; Enolase.
DR Pfam: PF00113; enolase; 1.
DR ProDom: PD000902; Enolase; 1.
FT NON_TER 1
SQ SEQUENCE 148 AA; 15554 MW; 3EF31039835D84A0 CRC64;

Query Match 42.6%; Score 52; DB 5; Length 148;
Best Local Similarity 58.8%; Pred. No. 4;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
Db 13 ANAILGVSLAVCKAGAA 29

```


GenCore version: 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:35:23 Search time 3.00429 seconds
(without alignments)
345.142 Million cell updates/sec

Title: US-09-727-739B-19

Perfect score: 122

Sequence: 1 MRVSIQHCALALLGLAALICSGAA 25

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	90.2	115	1	SMS2_ONCMY
2	74	60.7	114	1	SMSA_CARAU
3	73	59.8	120	1	SMS2_CARAU
4	64	52.5	114	1	SMS1_ICTPU
5	58	47.5	115	1	SMS1_RANRI
6	54	44.3	115	1	ENO_XENLA
7	54	44.3	815	1	AOXI_AERPE
8	53	43.4	116	1	SMS_CHICK
9	53	43.4	433	1	ENO_HOMGA
10	52	42.6	395	1	ENO_ALIMI
11	52	42.6	433	1	ENOA_ANAPL
12	52	42.6	433	1	ENOA_CHICK
13	52	42.6	433	1	ENOA_PYTRC
14	52	42.6	433	1	ENOA_SCEUN
15	52	42.6	433	1	ENOA_PRASC
16	52	42.6	433	1	ENOA_CHICK
17	52	42.6	433	1	ENOA_HUMAN
18	52	42.6	433	1	ENOA_MOUSE
19	52	42.6	433	1	ENOA_HUMAN
20	52	42.6	433	1	ENOA_RABIT
21	52	42.6	433	1	ENOA_RAT
22	52	42.6	433	1	ENOA_HUMAN
23	52	42.6	433	1	ENOA_MOUSE
24	52	42.6	433	1	ENOA_RAT
25	52	42.6	434	1	ENOA_CHICK
26	52	42.6	434	1	ENOA_LOLPE
27	52	42.6	434	1	ENO_SCHJA
28	52	42.6	434	1	ENO_SCHMA
29	50	41.0	439	1	ENO1_SCHPO
30	48	39.3	431	1	ENO_FASHE
31	48	39.3	433	1	ENOA_BOVIN
32	48	39.3	433	1	ENOA_HUMAN
33	48	39.3	433	1	ENOA_MOUSE
34	48	39.3	433	1	ENOA_RAT
35	47	38.5	276	1	UL34_HSV60
36	46	37.7	444	1	ENO1_HUMAN
37	46	37.7	444	1	ENO1_LYCES
38	46	37.7	702	1	SPE1_ORISA
39	45	36.9	419	1	KDAP_MOUSE
40	45	36.9	451	1	YOZ2_STRAT
41	44.5	36.5	116	1	SMS_HUMAN
42	44	36.1	239	1	CLDE_HUMAN
43	44	36.1	372	1	ENO_CHLRE
44	44	36.1	438	1	ENO_ASPOR
45	44	36.1	446	1	ENO_PLAFA

ALIGNMENTS

RESULT 1
SMS2_ONCMY STANDARD; PRT; 115 AA.
AC Q91194
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
[Tyr7,Gly10]somatostatin-14].
DE OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID:8022;
RN [1]
RP MEDLINE:95354921; PubMed:7628684;
RA Moore C.A., Kittilson J.D., Dahl S.K., Sheridan M.A.;
RT "Isolation and characterization of a cDNA encoding for
preprosomatostatin containing [Tyr7, Gly10]-somatostatin-14 from the
endocrine pancreas of rainbow trout, Oncorhynchus mykiss.";
RL Gen. Comp. Endocrinol. 98:253-261(1995).
CC 1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC -----
CC EMBL: U32471; AAC59695.1;
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin.1
KW Cleavage on pair of basic residues; Hormone; Signal; Multigene family.
FT SIGNAL
FT PROPEP 19 18
FT PEPTIDE 88 115 [Tyr21,Gly24]SOMATOSTATIN-28 (POTENTIAL).
FT PEPTIDE 102 115 [Tyr7,Gly10]SOMATOSTATIN-14.
FT DISULFID 104 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12963 MW; 520595023FCA6D91 CRC64;

Query Match 90.2%; Score 110; DB 1; Length 115;
Best Local Similarity 88.0%; Pred. No. 4.4e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRVSIQHCALALLGLAALICSGAA 25

Db 1 MKVCRHCALALLGLAALICSGAA 25

RESULT 2

SMSA_CARAU

```

ID SMSA_CARAU STANDARD; PRT; 114 AA.
AC Q9VGH5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin IA precursor [Contains: Somatostatin-26; Somatostatin-14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I and -II from goldfish brain.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)
CC EMBL; U00754; RAD09359.1;
CC EMBL; AF025686; AAF15306.1;
CC InterPro: IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues: Hormone; Signal; Multigene family.
CC SIGNAL 1 24 POTENTIAL.
CC PROPEP 25 88
CC PEPTIDE 89 114 SOMATOSTATIN-26 (POTENTIAL).
CC PEPTIDE 101 114 SOMATOSTATIN-14.
CC DISULFID 103 114 BY SIMILARITY.
CC SEQUENCE 114 AA; 13574 MW; B5920015E2D272A4 CRC64;
CC Query Match 60.7%; Score 74; DB 1; Length 114;
CC Best Local Similarity 64.0%; Pred. No. 0.00062;
CC Matches 16; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 MRVSIQHCALALLGLALAICSGAA 25
Db 1 MLSTRIQCALALLSLALAVCSAA 25
~SULT 3
~2_CARAU STANDARD; PRT; 120 AA.
AC Q9VGH4; Q9PTU2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28; [Tyr7,Gly10]somatostatin-14].
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Lin X.-W., Peter R.E.;
RT "Cloning and characterization of cDNAs encoding preprosomatostatin-I and -II from goldfish brain.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Otto C.J., Peter R.E.;
RT The expression of SRIF mRNA in the brain of goldfish.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)
CC EMBL; U00262; AAD09626.1;
CC EMBL; AF025686; AAF15306.1;
CC InterPro: IPR004250; Somatostatin.
CC Pfam; PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues: Hormone; Signal; Multigene family.
CC SIGNAL 1 23 POTENTIAL.
CC PROPEP 24 92
CC PEPTIDE 93 120 [Tyr21,Gly24]SOMATOSTATIN-28.
CC PEPTIDE 107 120 [Tyr7,Gly10]SOMATOSTATIN-14.
CC DISULFID 109 120 BY SIMILARITY.
CC CONFLICT 51 51 Q->RW (IN REF. 2).
CC SEQUENCE 120 AA; 13723 MW; 98957D68011A651A CRC64;
CC Query Match 59.8%; Score 73; DB 1; Length 120;
CC Best Local Similarity 54.2%; Pred. No. 0.0009;
CC Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRVSIQHCALALLGLALAICSGQA 24
Db 1 MRCELHLYALLGLSLVLCGRCA 24

RESULT 4
SMS1 ICTPU STANDARD; PRT; 114 AA.
AC P01171;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin I precursor [Contains: Somatostatin-14 (SS-14)].
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85303576; PubMed=2863931;
RA Dixon-J.E., Andrews P.C.;
RT "Somatostatin of the channel catfish.";
RL Adv. Exp. Med. Biol. 188:19-29(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=82263698; PubMed=6179939;
RA Minch C.D., Taylor W.L., Magazin M.D., Tavlanini M.A., Collier K.J.,
RA Weith H.L., Dixon J.E.;
RT "The structure of cloned DNA complementary to catfish pancreatic somatostatin-14 messenger RNA.";
RL J. Biol. Chem. 257:10372-10377(1982).
RN [3]
RP SEQUENCE OF 82-114 FROM N.A.
RC MEDLINE=82082515; PubMed=6171821;
RA Taylor W.L., Collier K.J., Deschenes R.J., Weith H.L., Dixon J.E.;
RT "Sequence analysis of a cDNA coding for a pancreatic precursor to somatostatin.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:6694-6698(1981).

```


[4]
RN SEQUENCE OF 101-114.
RA MEDLINE=81264223; PubMed=6114953;
RX Andrews P.C., Dixon J.E.;
RT "Isolation and structure of a peptide hormone predicted from a mRNA
sequence. A second somatostatin from the catfish pancreas.";
RL J. Biol. Chem. 256:8267-8270(1981).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PANCREAS.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M25903; AAA49339.1;
CC EMBL: V00607; CA23877.1;
CC EMBL: V00608; CA23878.1;
CC PIR: S00292; R11051.
CC InterPro: IPR004250; Somatostatin.
CC Pfam: PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Signal;
CC Multigene family.
CC SIGNAL 1 24 PROBABLE.
CC PEPTIDE 101 114 SOMATOSTATIN-14.
CC DISULFID 103 114
CC CONFLICT 62 62
CC SEQUENCE 114 AA; 12419 MW; FEE0F2C76F74D99F CRC64;
Query Match 52.5%; Score 64; DB 1; Length 114;
Best Local Similarity 52.0%; Pred. No. 0.017;
Matches 13; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
DY 1 MRYVQTHCALALGLALALICSGAA 25
DB 1 MPSTRIQCALALLAVALSVCSVSGA 25
RESULT 5
D SMSL_RANRI STANDARD; PRT; 115 AA.
AC P87384; Q9PSI8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin 1 precursor (PSS1) [Contains: Somatostatin-14 (S-I) (SSSI)].
DE Rana ridibunda (Laughing frog) (Marsh frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97057290; PubMed=8901629;
RA Tostivint H., Lihmann I., Buchard C., Vieau D., Coulouarn Y.,
RA Fournier A., Conlon J.M., Vaudry H.;
RT Occurrence of two somatostatin variants in the frog brain:
RT characterization of the cDNAs, distribution of the mRNAs, and
RT receptor-binding affinities of the peptides".
RL Proc. Natl. Acad. Sci. U.S.A. 93:12605-12610(1996).
RN [2]
RP SEQUENCE OF 102-115.
RC TISSUE=Brain;
RX MEDLINE=93038702; PubMed=1358069;
RA Vaudry H., Charrel N., Conlon J.M.;
RT "Isolation of [Pro2,Met13]somatostatin-14 and somatostatin-14 from the
RT frog brain reveals the existence of a somatostatin gene family in a

tetrapod.";
RL Biochem. Biophys. Res. Commun. 188:477-482(1992).
CC -1- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U68136; AAC60093.1;
CC InterPro: IPR004250; Somatostatin.
CC Pfam: PF03002; Somatostatin; 1.
CC Cleavage on pair of basic residues; Hormone; Multigene family; Signal.
CC SIGNAL 1 24 BY SIMILARITY.
CC PROPEP 25 99 BY SIMILARITY.
CC PEPTIDE 102 115 SOMATOSTATIN-14.
CC DISULFID 104 115 BY SIMILARITY.
CC SEQUENCE 115 AA; 12691 MW; 349756FEB4ABE213 CRC64;
Query Match 47.5%; Score 58; DB 1; Length 115;
Best Local Similarity 56.0%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
DY 1 MRYVQTHCALALGLALALICSGAA 25
DB 1 MOSCRVQCALTLALSALAINISAA 25
RESULT 6
ENO_XENLA STANDARD; PRT; 433 AA.
ID P08734;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).
GN ENO1
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88268812; PubMed=3390159;
RA Segil N., Shrutkowski A., Dworkin M.B., Dworkin-Rastl E.;
RT "Enolase isoenzymes in adult and developing Xenopus laevis and
RT characterization of a cloned enolase sequence".
RL Biochem. J. 251:31-39(1988).
CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate -> phosphoenolpyruvate +
CC H(2)O.
CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER.
CC -1- PATHWAY: Glycolysis.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y00718; CAA68706.1;
DR

DR PIR: S00463; NOXL
 DR HSSP: P56252; IPDZ
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF00113; enolase; 1.
 DR PRINTS: P00148; ENOLASE.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Lyase; Glycolysis; Magnesium.
 FT INIT_MET 0
 FT ACT_SITE 157 157
 FT METAL 244 244
 FT METAL 292 292
 FT METAL 317 317
 SQ SEQUENCE 433 AA; 47373 MW; 83C922489B530EB4 CRC64;

BY SIMILARITY.
 MAGNESIUM (BY SIMILARITY).
 MAGNESIUM (BY SIMILARITY).
 MAGNESIUM (BY SIMILARITY).
 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 47373 MW; 83C922489B530EB4 CRC64;

Query Match 44.3%; Score 54; DB 1; Length 433;
 Best Local Similarity 64.7%; Pred. No. 1.4;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 9 ALALLGLALALICSGAA 25
 | | | | | : | | | | |
 DB 107 ANALLGSLAVCKAGAA 123

RESULT 7
 ID AXIL_AERPE STANDARD; PRT; 815 AA.
 AC Q5IDX6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Heme-copper oxidase subunit I+III (EC 1.9.3.-).
 GN AOXB OR APOE793.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RA Wakagi T., Ishikawa R.;
 RT "Heme-copper-oxidase";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RA MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1";
 DNA Res. 6:83-101(1999).
 CC -I- COFACTOR: TWO HEME GROUPS AND COPPER B (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE HEME-COPPER
 CC RESPIRATORY OXIDASE FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE CYTOCHROME C
 CC OXIDASE SUBUNIT 3 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB020482; BA866072.1;

DR EMBL: AF000060; BAA79771.1; ALT_INIT.
 DR HSSP: P18401; IFFT.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: P00510; COX3; 1.
 DR ProDom: PD000382; Cytc_oxdse_III; 1.
 DR PROSITE: PS00077; COX1; 1.
 DR PROSITE: PS0253; COX3; 1.
 KW Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain;
 FT DOMAIN 1 467
 FT METAL 545 815
 FT METAL 70 70
 FT METAL 248 248
 FT METAL 252 252
 FT METAL 297 297
 FT METAL 298 298
 FT METAL 383 383
 FT METAL 385 385
 FT METAL 385 385
 FT METAL 26 46
 FT TRANSMEM 71 91
 FT TRANSMEM 105 125
 FT TRANSMEM 157 177
 FT TRANSMEM 197 217
 FT TRANSMEM 242 262
 FT TRANSMEM 281 301
 FT TRANSMEM 314 334
 FT TRANSMEM 339 359
 FT TRANSMEM 380 400
 FT TRANSMEM 419 439
 FT TRANSMEM 463 483
 FT TRANSMEM 580 600
 FT TRANSMEM 637 657
 FT TRANSMEM 683 703
 FT TRANSMEM 708 728
 FT TRANSMEM 736 756
 FT TRANSMEM 758 778
 FT TRANSMEM 791 811
 SQ SEQUENCE 815 AA; 90860 MW; 9E911C57BEF9BBE2 CRC64;

Query Match 44.3%; Score 54; DB 1; Length 815;
 Best Local Similarity 52.2%; Pred. No. 2.4;
 Matches 12; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 RVSOIHCALALIGLALALICSGA 24
 | | | | | : | | | | |
 DB 415 RLGRHFALMLGVALIFLQFA 437

RESULT 8
 ID SMS_CHICK STANDARD; PRT; 116 AA.
 AC F33094;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Somatostatin precursor [Contains: Somatostatin-28; Somatostatin-14].
 GN SST.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Naka K., Kobayashi T., Karahashi K., Kato S., Yamamoto H.,
 Yonekura H., Okamoto H.;
 RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
 CC -I- SUBCELLULAR LOCATION: Secreted.

CC -1- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X60191; CAA42747.1;
 DR PIR: S20630; S20630.
 DR InterPro: IPR004250; Somatostatin.
 DR Pfam: PF03002; Somatostatin; 1.
 KW Cleavage on pair of basic residues; Hormone; Signal.
 FT SIGNAL 1 24 BY SIMILARITY.
 FT PROPEP 25 88 BY SIMILARITY.
 FT T 89 116 SOMATOSTATIN-28.
 FT PEPTIDE 103 116 SOMATOSTATIN-14.
 FT DISULFID 105 116
 SQ SEQUENCE 116 AA; 12675 MW; 8A5B9BD8A291BA CRC64;
 Query Match 43.4%; Score 53; DB 1; Length 116;
 Best Local Similarity 52.4%; Pred. No. 0.63;
 Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 5 QIHCALGLALAIACSGAA 25
 Db 5 RLQCALALLSLAVGTVSAA 25
 RESULT 9
 ENO_HOMGA STANDARD; PRT; 433 AA.
 AC P56252; (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
 DE glyceralate hydro-lyase).
 OS Homarus gammarus (European lobster) (Homarus vulgaris).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 CC Astacodea; Nephropoidea; Nephropidae; Homarus.
 CX NCBI_TaxID=6707;
 N SEQUENCE AND X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RP TISSUE=Muscle.
 RX MEDLINE=96038209; PubMed=7547999;
 RA Duquerois S; Camus C; Janin J.;
 FT "X-ray structure and catalytic mechanism of lobster enolase."
 RL Biochemistry 34:12513-12523(1995).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
 CC H(2O).
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
 CC THE DIMER.
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
 DR PDB; 1PDY; 14-NOV-95.
 DR PDB; 1PDZ; 14-NOV-95.
 DR InterPro: IPR000941; Enolase.
 DR PRINTS: PR00148; Enolase.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Lyase; Glycolysis; Magnesium; Acetylation; 3D-structure.
 FT MOD_RES 157 157
 FT ACT_SITE 244 244
 FT METAL 294 294 MAGNESIUM (BY SIMILARITY)
 FT METAL 294 294 MAGNESIUM (BY SIMILARITY)

FT METAL 319 319 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 47041 MW; CC7DA44B67E9DE3D CRC64;
 Query Match 43.4%; Score 53; DB 1; Length 433;
 Best Local Similarity 64.7%; Pred. No. 1.9;
 Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 9 ALALLGLALAICSGAA 25
 Db 107 ANAILGLSLAICAKAGAA 123
 RESULT 10
 ENO_ALLMI STANDARD; PRT; 395 AA.
 AC P42897;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
 DE glyceralate hydro-lyase) (Fragment)
 OS Alligator mississippiensis (American alligator)
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Crocodylia; Alligatorinae; Alligator.
 CX NCBI_TaxID=8496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94195794; PubMed=8146164;
 RA Hedges S.B.;
 FT "Molecular evidence for the origin of birds."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2621-2624(1994).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
 CC H(2O).
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
 CC THE DIMER (BY SIMILARITY).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L28078; AAA53671.1;
 DR HSP; P56252; IPDZ.
 DR InterPro: IPR000941; Enolase.
 DR Pfam: PF00113; enolase; 1.
 DR ProDom: PD000902; Enolase; 1.
 DR TIGRFAMs: TIGR01060; eno; 1.
 DR PROSITE: PS00164; ENOLASE; 1.
 KW Lyase; Glycolysis; Magnesium.
 FT NON_TER 1 1
 FT ACT_SITE 136 136 BY SIMILARITY.
 FT METAL 223 223 MAGNESIUM (BY SIMILARITY).
 FT METAL 271 271 MAGNESIUM (BY SIMILARITY).
 FT METAL 296 296 MAGNESIUM (BY SIMILARITY).
 FT NON_TER 395 395
 SQ SEQUENCE 395 AA; 42884 MW; B43E91228E9110B5 CRC64;
 Query Match 42.6%; Score 52; DB 1; Length 395;
 Best Local Similarity 58.8%; Pred. No. 2.5;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 9 ALALLGLALAICSGAA 25
 Db 86 ANAILGLSLAICAKAGAA 102

```

RESULT 11
ENOA_ALLMI STANDARD; PRT; 433 AA.
AC Q2PVK2:2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
DE (Phosphopyruvate hydratase).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylia; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic lens;
MEDLINE=89079778; PubMed=2462567;
RA Wislow G.J., Lietman T., Williams L.A., Stapel S.O., de Jong W.W.,
RA Horwitz J., Piatigorsky J.;
RT "Tau-crystallin/alpha-enolase: one gene encodes both an enzyme and a
RT lens structural protein."
RL J. Cell Biol. 107:2729-2736(1988).
CC - FUNCTION: BOTH AN ENZYME AND A LENS STRUCTURAL PROTEIN.
CC - CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
CC H(2O).
CC - COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER.
CC - PATHWAY: Glycolysis.
CC - SUBUNIT: HOMODIMER.
CC - SUBCELLULAR LOCATION: Cytoplasmic.
CC - SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20749; AAA49218.1;
CC EMBL; X14195; CAA32409.1;
CC PIR; A32132; A32132.
CC HSSP; P56252; 1PD2.
CC InterPro; IPR000941; Enolase.
CC Pfam; PF00113; enolase; 1.
CC PRINTS; PR00148; ENOLASE.
CC PRODOM; PD000902; Enolase; 1.
CC TIGRFAMs; TIGR01060; eno; 1.
CC PROSITE; PS00164; ENOLASE; 1.
CC PROSITE; TIGR01060; eno; 1.
CC Lyase; Glycolysis; Magnesium; Eye lens protein.
KW INIT_MET 0
FT ACT_SITE 157 157 BY SIMILARITY.
FT METAL 244 244 MAGNESIUM (BY SIMILARITY).
FT METAL 292 292 MAGNESIUM (BY SIMILARITY).
FT METAL 317 317 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 433 AA; 47108 MW; 1AED78B08A66E84D CRC64;
Query Match 42.6%; Score 52; DB 1; Length 433;
Best Local Similarity 58.8%; Pred. No. 2.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 9 ALALGLALALICSGAA 25
DB 107 ANAILGVSLAVCKAGAA 123
| : : : : : |
| : : : : : |

RESULT 13
ENOA_CHICK STANDARD; PRT; 433 AA.
ID ENOA_CHICK
AC P51913;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
DE (Phosphopyruvate hydratase).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; Tissue=kidney;
MEDLINE=95355305; PubMed=7629021;
RX

```

RA Tanaka M., Maeda K., Nakashima K.;
 PT "chicken alpha-enolase but not beta-enolase has a Src-dependent
 RT tyrosine phosphorylation site: cDNA cloning and nucleotide sequence
 RL J. Biol. Chem., 117:554-559(1995).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
 CC H₂O
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
 CC THE DIMER (BY SIMILARITY).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D37900; BAA07132.1;
 CC HSSP: P56252; IPDZ.
 CC InterPro: IPR000941; Enolase.
 CC Pfam: PF00113; enolase; 1.
 CC PRINTS: PR00148; ENOLASE.
 CC PRODOM: PD000902; Enolase; 1.
 CC TIGRFAMS: TIGR01060; eno; 1.
 CC PROSITE: PS00164; ENOLASE; 1.
 CC Lyase; Glycolysis; Magnesium; Multigene family.
 CC INIT_MET 0
 CC ACT_SITE 157 157 BY SIMILARITY.
 CC METAL 244 244 MAGNESIUM (BY SIMILARITY).
 CC METAL 292 292 MAGNESIUM (BY SIMILARITY).
 CC METAL 317 317 MAGNESIUM (BY SIMILARITY).
 CC SEQUENCE 433 AA; 47173 MW; 93AD6B0A/AD99510 CRC64;
 CC
 CC Query Match 42.6%; Score 52; DB 1; Length 433;
 CC Best Local Similarity 58.8%; Pred. No. 2.7; Mismatches 4; Indels 0; Gaps 0;
 CC Matches 10; Conservative 4;
 CC
 CC QY 9 ALALLGLALAICSGAA 25
 CC Db 107 ANAILGVSLAVCKAGAA 123

RESULT 14
 ENOA_PYTRG
 ID ENOA_PYTRG STANDARD; PRT; 433 AA.
 AC Q9W7L0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
 DE (Phosphopyruvate hydratase).
 OS Python regius (Ball python).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Henophidia;
 OC Pythonidae; Python.
 CC NCBI_TaxID=51751;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Muscle;
 CC MEDLINE=99439677; PubMed=10508547;
 CC Manner H., Li S.-L.;
 CC "Molecular evidence for a clade of turtles."
 CC Mol. Phylogenet. Evol. 13:144-148(1999).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
 CC H₂O.
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
 CC THE DIMER (BY SIMILARITY).
 CC -1- PATHWAY: Glycolysis.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF072589; AAD41646.1;
 CC HSSP: P56252; IPDZ.
 CC InterPro: IPR000941; Enolase.
 CC Pfam: PF00113; enolase; 1.
 CC PRINTS: PR00148; ENOLASE.
 CC PRODOM: PD000902; Enolase; 1.
 CC TIGRFAMS: TIGR01060; eno; 1.
 CC PROSITE: PS00164; ENOLASE; 1.
 CC Lyase; Glycolysis; Magnesium; Multigene family.
 CC INIT_MET 0
 CC ACT_SITE 244 244 BY SIMILARITY.
 CC METAL 292 292 MAGNESIUM (BY SIMILARITY).
 CC METAL 317 317 MAGNESIUM (BY SIMILARITY).
 CC SEQUENCE 433 AA; 47439 MW; B0AD198562629D23 CRC64;
 CC
 CC Query Match 42.6%; Score 52; DB 1; Length 433;
 CC Best Local Similarity 58.8%; Pred. No. 2.7; Mismatches 4; Indels 0; Gaps 0;
 CC Matches 10; Conservative 4;
 CC
 CC QY 9 ALALLGLALAICSGAA 25
 CC Db 107 ANAILGVSLAVCKAGAA 123

RESULT 15
 ENOA_SCEUN
 ID ENOA_SCEUN STANDARD; PRT; 433 AA.
 AC Q9W7L2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Alpha enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase)
 DE (Phosphopyruvate hydratase).
 OS Sceloporus undulatus (Eastern fence lizard) (Skink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;
 OC Sceloporus.
 CC NCBI_TaxID=8520;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Muscle;
 CC MEDLINE=99439677; PubMed=10508547;
 CC Manner H., Li S.-L.;
 CC "Molecular evidence for a clade of turtles."
 CC Mol. Phylogenet. Evol. 13:144-148(1999).
 CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate - phosphoenolpyruvate +
 CC H₂O.
 CC -1- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
 CC THE DIMER (BY SIMILARITY).
 CC -1- PATHWAY: Glycolysis.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

```
CC      EMBL; AF072587; AAD41644.1; -  
DR     HSSP; ES6252; IPDZ;  
DR     InterPro; IPR000941; Enlase.  
DR     Pfam; PF00113; enlase; 2.  
DR     PRINTS; PR00148; ENOLASE..  
DR     ProDom; PD000902; Enlase.. 1  
DR     TIGRFAMs; TIGR01060; eno: 1.  
DR     PROSITE; PS00164; ENOLASE; 1.  
KW     Lyase; Glycylis; Magnesium;  
FT     INIT_MET    0  
FF     METAL       244  
FT     METAL       244  
FF     METAL       292  
FT     METAL       317  
SQ     SEQUENCE   433 AA;  47362 MW;  7CD89EB40529301 CRC64;  
  
                BY SIMILARITY.  
            MAGNESIUM (BY SIMILARITY).  
            MAGNESIUM (BY SIMILARITY).  
            MAGNESIUM (BY SIMILARITY).  
            MAGNESIUM (BY SIMILARITY).  
            MAGNESIUM (BY SIMILARITY);  
            7CD89EB40529301 CRC64;  
  
Query Match          42.6%; Score 52; DB 1; Length 433;  
Best Local Similarity 58.8%. Pred. No. 2.7;  
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
OY      9 ALALLGLALAICSGAA 25  
         |::|::||:  
nb     107 ANALGVSLAVCKAGAA 123
```

Search completed: March 21, 2003, 11:38:45
Job time : 5.00429 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:38:56 ; Search time 5.36481 Seconds
(without alignments)
447.986 Million cell updates/sec

Title: US-09-727-739b-19
Perfect score: 122
Sequence: 1 MRVSOIHCALLGLALAIQSGAA 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	90.2	115	2 I51064	somatostatin II pr
2	64	52.5	114	1 RIIDS1	somatostatin-14 pr
3	64	52.5	114	2 I50798	preprosomatostatin
4	58	47.5	115	2 JC6166	somatostatin-14 pr
5	54	44.3	434	1 NOXL	phosphopyruvate hy
6	54	44.3	818	2 C72671	probable cytochrom
7	53	43.4	116	1 S20630	somatostatin-precu
8	52	42.6	395	2 I50026	phosphopyruvate hy
9	52	42.6	431	2 A33665	phosphopyruvate hy
10	52	42.6	433	2 A37210	phosphopyruvate hy
11	52	42.6	434	1 NQHSB	phosphopyruvate hy
12	52	42.6	434	2 A24742	phosphopyruvate hy
13	52	42.6	434	2 S10247	phosphopyruvate hy
14	52	42.6	434	2 A32132	phosphopyruvate hy
15	52	42.6	434	2 S02072	phosphopyruvate hy
16	52	42.6	434	2 S06756	phosphopyruvate hy
17	52	42.6	434	2 JC4186	phosphopyruvate hy
18	52	42.6	434	2 JC4187	phosphopyruvate hy
19	52	42.6	434	2 JC1039	phosphopyruvate hy
20	52	42.6	434	2 F83244	hypothetical prote
21	51	41.8	120	2 F83264	hypothetical prote
22	50.5	41.4	948	2 D69819	3-hydroxybutyryl-Co
23	50	41.0	275	2 JC4036	phosphopyruvate hy
24	50	41.0	439	2 T39737	enolase - fission
25	50	41.0	1582	2 T15308	hypothetical prote
26	50	41.0	531	2 B83082	probable binding p
27	49	40.2	531	2 H87617	hypothetical prote
28	48	39.3	174	2 H87617	phosphopyruvate hy
29	48	39.3	434	2 S10246	phosphopyruvate hy

30	48	39.3	434	2 A29170	phosphopyruvate hy
31	48	39.3	434	2 A23126	phosphopyruvate hy
32	47	38.5	106	2 F86565	CT483 hypothetical
33	47	38.5	106	2 C72057	conserved hypothet
34	47	38.5	181	2 C81041	pilus assembly pro
35	47	38.5	279	2 T09328	probable membrane
36	47	38.5	458	2 I37360	phosphopyruvate hy
37	47	38.5	495	2 C71410	hypothetical prote
38	46	37.7	316	2 A83588	hypothetical prote
39	46	37.7	326	2 J01186	phosphopyruvate hy
40	46	37.7	399	2 G98197	hypothetical prote
41	46	37.7	399	2 AC3089	conserved hypothet
42	46	37.7	444	2 J01185	phosphopyruvate hy
43	46	37.7	849	2 D82785	mannosyltransferas
44	45.5	37.3	439	2 D83345	L-sorbose dehydr
45	45	36.9	142	2 H83450	hypothetical prote

ALIGNMENTS

RESULT 1
I51064
somatostatin II precursor - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51064
R:Moore, C.A.; Kittilson, J.D.; Dahl, S.K.; Sheridan, M.A.
Gen. Comp. Endocrinol. 98, 253-261, 1995
A:Title: Isolation and characterization of a cDNA encoding for preprosomatostatin c
A:Reference number: I51064; MUID:95354921; PMID:7628684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-115 <MOO>
A:Cross-references: EMBL:U32471; NID:g975344; PIDN:AAC59695.1; PID:g975345
C:Superfamily: Somatostatin
Query Match 90.2%; Score 110; DB 2; Length 115;
Best Local Similarity 88.0%; Pred. No. 8.1e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRVSOIHCALLGLALAIQSGAA 25
DB 1 MKVCRHICALALLGLALAIQSGAA 25
RESULT 2
RIIDS1
somatostatin-14 precursor - channel catfish
N:Alternate names: somatostatin I
N:Contains: somatostatin-14
C:Species: Ictalurus punctatus (channel catfish)
C>Date: 30-Jun-1980 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C:Accession: S00292; A93897; A2334; A01435
R:Mintz, C.D.; Taylor, W.L.; Magazin, M.; Tavariani, M.A.; Collier, K.; Weith, H.L.
J. Biol. Chem. 257, 10372-10377, 1982
A:Title: The structure of cloned DNA complementary to catfish pancreatic somatostat
A:Reference number: S00292; MUID:82265698; PMID:6179939
A:Accession: S00292
A:Molecule type: mRNA
A:Residues: 1-114 <MIN>
A:Cross-references: EMBL:V00607; NID:g64017; PIDN:CAA23877.1; PID:g64018
R:Taylor, W.L.; Collier, K.J.; Deschenes, R.J.; Weith, H.L.; Dixon, J.E.
Proc. Natl. Acad. Sci. U.S.A. 78, 6694-6698, 1981
A:Title: Sequence analysis of a cDNA coding for a pancreatic precursor to somatosta
A:Reference number: A93897; MUID:82082515; PMID:6171821
A:Accession: A93897
A:Molecule type: mRNA
A:Residues: 82-108 <TAG>
A:Cross-references: GB:J00944
R:Andrews, P.C.; Dixon, J.E.
J. Biol. Chem. 256, 8267-8270, 1981

Db 418 RLGIHFALMLGVALTFLPOFA 440

RESULT 7

S20630
somatostatin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S20630
R:Nata, K.; Kobayashi, T.; Karahashi, K.; Kato, S.; Yamamoto, H.; Yonekura, H.; Okamoto, submitted to the EMBL Data Library, June 1991
A:Description: Nucleotide sequence determination of chicken somatostatin precursor cDNA.
A:Reference number: S20630
A:Accession: S20630
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-116 <NAT>
A:Cross-references: EMBL:X60191; NID:g62985; PIDN:CAA42747.1; PID:g62986
C:Superfamily: somatostatin

Query Match 43.4%; Score 53; DB 1; Length 116;
Best Local Similarity 52.4%; Pred. No. 1.4;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 QHICALALLGLAALICSGGAA 25
DB 5 RLQCALALLSIALAVGTSSAA 25

RESULT 8

I50026
phosphorylase hydratase (EC 4.2.1.11) alpha - American alligator (fragment)
N:Alternate names: alpha-enolase
C:Species: Alligator mississippiensis (American alligator)
C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 22-Jun-1999
C:Accession: I50026
R:Hedges, S.B.
Proc Natl. Acad. Sci. U.S.A. 91, 2621-2624, 1994
A:Title: Molecular evidence for the origin of birds
A:Reference number: A53470; MUID:94195794; PMID:8146164
A:Accession: I50026
A>Status: preliminary; nucleic acid sequence not shown; translation not shown; translated
A:Molecule type: mRNA
A:Residues: 1-395 <HED>
A:Cross-references: GB:L28078; NID:g472796; PIDN:AAA53671.1; PID:g472797
C:Superfamily: enolase
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 395;
Best Local Similarity 58.8%; Pred. No. 5.3;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGGAA 25
DB 86 ANAILGVSLAVCKAGAA 102

RESULT 9

A53665
phosphorylase hydratase (EC 4.2.1.11) - liver fluke
C:Species: Fasciola hepatica (liver fluke)
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 22-Jun-1999
C:Accession: A53665
R:David, R.E.; Singh, H.; Botka, C.; Hardwick, C.; Ashraf el Meanawy, M.; Villanueva, J. J. Biol. Chem. 269, 20026-20030, 1994
A:Title: RNA trans-splicing in Fasciola hepatica. Identification of a spliced leader (SL) in the 5' non-coding region of the rRNA gene.
A:Reference number: A53665; MUID:94327554; PMID:8051087
A:Accession: A53665
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-431 <DAV>
A:Cross-references: GB:U10297; NID:g499266; PIDN:AAA57450.1; PID:g499267
C:Superfamily: enolase

C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 431;
Best Local Similarity 58.8%; Pred. No. 5.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGGAA 25
DB 107 ANAILGVSLAVCKAGAA 123

RESULT 10

A37210
phosphopyruvate hydratase (EC 4.2.1.11) beta - rabbit
N:Alternate names: enolase beta
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 05-Aug-1994
C:Accession: A37210
R:Chin, C.C.Q.

J. Protein Chem. 9, 427-432, 1990
A:Title: The primary structure of rabbit muscle enolase.
A:Reference number: A37210; MUID:91113295; PMID:2275753
A:Accession: A37210
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-433 <CHI>
C:Superfamily: enolase
C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase
F:1/Modified site: acetylated amino end (Ala) #status experimental

Query Match 42.6%; Score 52; DB 2; Length 433;
Best Local Similarity 58.8%; Pred. No. 5.7;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSGGAA 25
DB 107 ANAILGVSLAVCKAGAA 123

RESULT 11

NOHUG
phosphopyruvate hydratase (EC 4.2.1.11) gamma - human
N:Alternate names: enolase gamma; neuron-specific enolase
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Jun-1999
C:Accession: JU0060; S16163; S02077; I56569; S02616; S38303
R:Oliva, D.; Barba, G.; Barbieri, G.; Giallongo, A.; Feo, S.
Gene 79, 355-360, 1989
A:Title: Cloning, expression and sequence homologies of cDNA for human gamma enolase
A:Reference number: JU0060; MUID:90006764; PMID:2792767
A:Accession: JU0060
A:Molecule type: mRNA
A:Residues: 1-434 <OLI>
A:Cross-references: GB:M22349; NID:g951199; PIDN:AAB59554.1; PID:g182116; GB:M27833
R:Oliva, D.; Cali, L.; Feo, S.; Giallongo, A.
Genomics 10, 157-165, 1991
A:Title: Complete structure of the human gene encoding neuron-specific enolase.
A:Reference number: S16163; MUID:91257823; PMID:2045099
A:Accession: S16163
A:Molecule type: DNA
A:Residues: 1-434 <OLI>
A:Cross-references: GB:X51956; NID:g31164; PIDN:CAA36215.1; PID:g31165
R:McAleese, S.M.; Dunbar, B.; Fothergill, J.E.; Hinks, L.J.; Day, I.N.M.
Eur. J. Biochem. 178, 413-417, 1988
A:Title: Complete amino acid sequence of the neurone-specific gamma isozyme of enol
A:Reference number: S02077; MUID:89091176; PMID:3208766
A:Accession: S02077
A:Molecule type: mRNA
A:Residues: 2-3, 'O', 5-239, 'M', 241-434 <MCA>
A:Cross-references: EMBL:X13120; NID:g31145; PIDN:CAA31512.1; PID:g930063
A:Note: Part of this sequence was confirmed by protein sequencing
A:Note: 264-Ala and 395-Ala were also found
R:Van Obberghen, E.; Kamholz, J.; Bishop, J.G.

RESULT 14

S10247
 phosphopyruvate hydratase (EC 4.2.1.11) gamma - mouse
 N:Alternate names: 2-phosphoglycerate dehydratase; enolase gamma
 C:Species: Mus musculus (house mouse)
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 22-Jun-1999
 C:Accession: S10247
 R:Raghad, M.; Dumont, X.; Chalon, P.; Lelias, J.M.; Lamande, N.; Lucas, M.; Lazar, M.; O
 Nucleic Acids Res 18, 3638, 1990
 A:Title: Nucleotide sequences of cDNAs alpha and gamma enolase mRNAs from mouse brain.
 A:Reference number: S10246; MUID:90301487; PMID:2362815
 A:Accession: S10247
 A:Molecule type: mRNA
 A:Residues: 1-434 <KAG>
 A:Cross-references: EMBL:X52380; NID:g55494; PIDN:CAA36606.1; PID:g55495
 C:Superfamily: enolase
 C:Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 434;
 Best Local Similarity 58.8%; Pred. No. 5.7;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALATCSOGAA 25
 | : | | : | | : | |
 Db 108 ANAILGVSLAVCKAGAA 124

RESULT 15

A32132
 phosphopyruvate hydratase (EC 4.2.1.11) alpha - duck
 N:Alternate names: enolase alpha; tau-crystallin
 C:Species: Anas platyrhynchos (domestic duck)
 C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 21-Jul-2000
 C:Accession: A32132
 R:Wistow, G.J.; Lietman, T.; Williams, L.A.; Stapel, S.O.; de Jong, W.W.; Horwitz, J.; P
 J. Cell Biol. 107, 2729-2736, 1988
 A:Title: Tau-crystallin/alpha-enolase: one gene encodes both an enzyme and a lens struct
 A:Reference number: A92750; MUID:89079778; PMID:2462567
 A:Accession: A32132
 A:Molecule type: mRNA
 A:Residues: 1-434 <WIS>
 A:Cross-references: GB:X14195; NID:g62455; PIDN:CAA32409.1; PID:g62456
 C:Superfamily: enolase
 C:Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase

Query Match 42.6%; Score 52; DB 2; Length 434;
 Best Local Similarity 58.8%; Pred. No. 5.7;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALATCSOGAA 25
 | : | | : | | : | |
 Db 108 ANAILGVSLAVCKAGAA 124

Search completed: March 21, 2003, 11:41:45
 Job time: 7.36481 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 11:53:33 ; Search time 4.82833 Seconds
(without alignments)
276.816 Million cell updates/sec

Title: US-09-727-739B-19

Perfect score: 122

Sequence: 1 MRVSIHICALALLGLAICSGAA 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 221153 segs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2.6/ptodata/1/pubpaa/RCUS_PUBCOMB pep.*
- 8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 10: /cgn2.6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2.6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	41.0	1582	9 US-09-966-422B-11	Sequence 11, Appl
2	48	39.3	120	9 US-09-796-692-2277	Sequence 2277, Ap
3	48	39.3	162	9 US-09-796-692-2342	Sequence 2342, Ap
4	48	39.3	433	9 US-09-935-642-9	Sequence 9, Appl
5	48	39.3	1771	9 US-10-184-644-17	Sequence 17, Appl
6	47	38.5	423	10 US-09-843-164-10	Sequence 10, Appl
7	47	38.5	481	10 US-09-843-164-8	Sequence 8, Appl
8	47	38.5	502	10 US-09-843-164-14	Sequence 14, Appl
9	47	38.5	560	9 US-09-966-422B-2	Sequence 2, Appl
10	47	38.5	560	10 US-09-843-164-12	Sequence 12, Appl
11	47	38.5	769	10 US-09-843-164-2	Sequence 2, Appl
12	47	38.5	848	10 US-09-843-164-6	Sequence 6, Appl
13	46	37.7	400	9 US-09-971-228-12	Sequence 12, Appl
14	46	37.7	400	10 US-09-812-272-2	Sequence 2, Appl
15	46	37.7	400	12 US-10-037-616-27	Sequence 27, Appl
16	45	36.9	97	9 US-10-057-275-5	Sequence 5, Appl
17	45	36.9	97	9 US-09-920-137A-5	Sequence 5, Appl
18	45	36.9	419	10 US-09-789-919-44	Sequence 44, Appl
19	45	36.9	2694	9 US-10-184-644-207	Sequence 207, App

ALIGNMENTS

RESULT 1

US-09-966-422B-11
; Sequence 11, Application US/09966422B
; Publication No. US20030044892A1
; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBRM6, EXPRESSE
; FILE REFERENCE: D0040NF/3053-4119US3
; CURRENT APPLICATION NUMBER: US/09/966,422B
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 60/235,602
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/306,604
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/315,412
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1582
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

US-09-966-422B-11

Query Match 41.0%; Score 50; DB 9; Length 1582;

Best Local Similarity 50.0%; Pred No. 51;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 VSOIHCALALLGLAIC 20

Db 910 VSTIGCALISVCLALSVC 927

RESULT 2

US-09-796-692-2277
; Sequence 2277, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,779
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2277
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(120)
; OTHER INFORMATION: Xaa - Any amino acid
US-09-796-692-2277

Query Match 39.3%; Score 48; DB 9; Length 120;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24
| : : : : | : : |
Db 65 ANAILGSLAVCKAGA 80

RESULT 3
-09-796-692-2342
Sequence 2342, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Hannion, Jane
APPLICANT: Algate, Paul A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796.692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2277
LENGTH: 120
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(120)
OTHER INFORMATION: Xaa - Any amino acid
US-09-796-692-2277

Query Match 39.3%; Score 48; DB 9; Length 120;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24
| : : : : | : : |
Db 65 ANAILGSLAVCKAGA 80

RESULT 4
US-09-935-642-9
Sequence 9, Application US/09935642
Publication No. US20030044795A1
GENERAL INFORMATION:
APPLICANT: BYRJALSEN, Inger
APPLICANT: LARSEN, Peter
APPLICANT: STEPHEN, John
TITLE OF INVENTION: Biochemical Markers for the Human
TITLE OF INVENTION: Endometrium
FILE REFERENCE: 8969-014
CURRENT APPLICATION NUMBER: US/09/935,642
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: PCT/GB97/02394
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: PCT/GB9707132.8
PRIOR FILING DATE: 1997-04-08
PRIOR APPLICATION NUMBER: PCT/GB9618600.2
PRIOR FILING DATE: 1996-09-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 433
TYPE: PRT
ORGANISM: Homo sapiens
US-09-935-642-9

Query Match 39.3%; Score 48; DB 9; Length 433;
Best Local Similarity 56.2%; Pred. No. 27;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALALLGLALAICSQGA 24
| : : : : | : : |
Db 107 ANAILGSLAVCKAGA 122

RESULT 5
US-10-184-644-17
```

Sequence 17, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
PRIOR FILING DATE: 2002-06-28
Prior Application removed - See file Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 17
LENGTH: 1771
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-17

Query Match 39.3%; Score 48; DB 9; Length 1771;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 8 CALALLGLAALICSQGAA 25
DB 1452 CACAATGAATCATGAA 1469

RESULT 6
US-09-843-164-10
Sequence 10, Application US/09843164
Patent No. US20020061556A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
TITLE OF INVENTION: NO. US20020061556A1 Human Membrane Proteins and Polynucleotides
FILE REFERENCE: 07705.0014-00000
CURRENT APPLICATION NUMBER: US/09/843,164
CURRENT FILING DATE: 2001-04-27
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/199,950
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 423
TYPE: PRT
ORGANISM: homo sapiens
US-09-843-164-10

Query Match 38.5%; Score 47; DB 10; Length 423;
Best Local Similarity 45.0%; Pred. No. 37;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHCAALGLALALICSQ 22
DB 142 LSNVGCALSVTGLALTIVIFQ 161

RESULT 7
US-09-843-164-8
Sequence 8, Application US/09843164
Patent No. US20020061556A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade

APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
TITLE OF INVENTION: NO. US20020061556A1 Human Membrane Proteins and Polynucle
FILE REFERENCE: 07705.0014-00000
CURRENT APPLICATION NUMBER: US/09/843,164
CURRENT FILING DATE: 2001-04-27
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/199,950
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 481
TYPE: PRT
ORGANISM: homo sapiens
US-09-843-164-8

Query Match 38.5%; Score 47; DB 10; Length 481;
Best Local Similarity 45.0%; Pred. No. 42;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHCAALGLALALICSQ 22
DB 200 LSNVGCALSVTGLALTIVIFQ 219

RESULT 8
US-09-843-164-14
Sequence 14, Application US/09843164
Patent No. US20020061556A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wang, Xiaoming
APPLICANT: Scoville, John
TITLE OF INVENTION: NO. US20020061556A1 Human Membrane Proteins and Polynucle
FILE REFERENCE: 07705.0014-00000
CURRENT APPLICATION NUMBER: US/09/843,164
CURRENT FILING DATE: 2001-04-27
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/199,950
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 502
TYPE: PRT
ORGANISM: homo sapiens
US-09-843-164-14

Query Match 38.5%; Score 47; DB 10; Length 502;
Best Local Similarity 45.0%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSQIHCAALGLALALICSQ 22
DB 142 LSNVGCALSVTGLALTIVIFQ 161

RESULT 9
US-09-966-422B-2
Sequence 2, Application US/09966422B
Publication No. US20030044892A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY6, EXPRESSE
FILE REFERENCE: D0040NP/3053-4119S3
CURRENT APPLICATION NUMBER: US/09/966,422B
CURRENT FILING DATE: 2002-05-07
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,602
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/306,604
PRIOR FILING DATE: 2001-08-28

```

; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 560
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-966-422B-2

```

Query Match 38.5%; Score 47; DB 9; Length 560;
Best Local Similarity 45.0%; Pred. NO. 48;
Matches 9; Conservative 5; Mismatches 6; Indels

QY 3 VSQIHCALALGLALAICSQ 22
:|:|:|:|:|:|:|:|:|
Db 200 LSNVGCALSVTGLALTVEQ 219

RESULT 10
US-09-843-164-12
; Sequence 12, Application US/09843164
; Patent No. US20020061556A1

Query Match 38.5%; Score 47; DB 10; Length 560;
Best Local Similarity 45.0%; Pred. No. 48;
Matches 9: Conservative 5; Mismatches 6; Indels

QY 3 VSQIHCALALGLALACSQ 22
: | : | | : | | : |
Db 200 LSNVGCALSVTGLALTVEQ 219

RESULT 11
US-09-843-164-2
: Sequence 2, Application US/09843164
Patent No. US20020061556A1

Query Match 38.5%; Score 47; DB 10; Length 769;
Best Local Similarity 45.0%; Pred. No. 66;
Matches 9: Conservative 5; Mismatches 6; Indels

Qy 3 VSQIHCALALGLALACSQ 22
Db 488 LSNVGCALSVTGLALTVIFQ 507

RESULT 12
US-09-843-164-6
; Sequence 6, Application US/09843164
; Patent No. US20020061556A1

Query Match 38.5%; Score 47; DB 10; Length 848;
Best Local Similarity 45.0%; Pred. No. 72;
Matches 9; Conservative 5; Mismatches 6; Indels

QY . 3 VSQIHICALLGLALAICSQ 22
: | : | | | : | | | : |
Db . 488 LSNVGCALSVTGLALTVEFQ 507

RESULT 13
US-09-971-228-12
; Sequence 12, Application US/09971228
; Patent No. US20020155512A1
; GENERAL INFORMATION:
; APPLICANT: Liao, X. Charlene
; APPLICANT: Masuda, Esteban
; APPLICANT: Chu, Peter
; APPLICANT: Pardo, Jorge
; APPLICANT: Li, Congfen
; APPLICANT: Zhao, Haoran
; APPLICANT: Jiang, Yingping
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: EDG: Modulators of Lymphocyte Activation and Migration

Query Match 37.7%; Score 46; DB 9; Length 400;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 8: Conservative 3; Mismatches 5; Indels

QY . . . 6 IHCA L L L G L A I C S 21
: | | | : | | |

Db 195 LFCVLAFLGLAICA 210

RESULT 14

US-09-812-272-2
 ; Sequence 2, Application US/09812272
 ; Patent No. US20020045214A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pamela Lane
 ; APPLICANT: Ping Tsui
 ; APPLICANT: Nabil A. Elshourbagy
 ; TITLE OF INVENTION: RAT G-PROTEIN COUPLED RECEPTOR AXOR29
 ; FILE REFERENCE: GP-70685
 ; CURRENT APPLICATION NUMBER: US/09/812,272
 ; CURRENT FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: 60/191,153
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 400
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 US-09-812-272-2

Query Match 37.7%; Score 46; DB 10; Length 400;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 6 IHCAALLGLAICA 21

Db 195 LFCVLAFLGLAICA 210

RESULT 15

US-10-037-616-27
 ; Sequence 27, Application US/10037616
 ; Patent No. US20020123148A1
 ; GENERAL INFORMATION:
 ; APPLICANT: English, Denis
 ; APPLICANT: Kovacs, Richard J.
 ; APPLICANT: Rizzo, Maria T.
 ; APPLICANT: Sliva, Daniel T.
 ; TITLE OF INVENTION: Sphingolipid Compositions and Methods for Their Therapeutic Use
 ; FILE REFERENCE: 7042-119
 ; CURRENT APPLICATION NUMBER: US/10/037,616
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: US 60/243,887
 ; PRIOR FILING DATE: 2000-10-27
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 27
 ; LENGTH: 400
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-037-616-27

Query Match 37.7%; Score 46; DB 12; Length 400;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 6 IHCAALLGLAICA 21

Db 195 LFCVLAFLGLAICA 210

Search completed: March 21, 2003, 11:56:32
 Job time : 5.82833 secs

ult	Score	Query	Match	Length	DB	ID	Description
1	46	37.7	433	4	US-09-705-448-3		Sequence 3, Appli
2	46	37.7	445	4	US-08-974-691-6		Sequence 6, Appli
3	46	37.7	451	4	US-08-974-691-2		Sequence 2, Appli
4	45	36.9	419	4	US-08-974-691-3		Sequence 3, Appli
5	45	36.9	419	4	US-09-705-448-10		Sequence 10, Appl
6	44	36.1	350	4	US-08-637-670-25		Sequence 25, Appl
7	43	35.2	249	4	US-09-105-343A-2		Sequence 2, Appli
8	42	34.4	81	1	US-08-469-667-14		Sequence 14, Appl
9	42	34.4	81	1	US-09-224-110-14		Sequence 14, Appl
10	42	34.4	81	5	PCV-US95-07289-14		Sequence 14, Appl
11	42	34.4	395	1	US-08-723-938-3		Sequence 3, Appli
12	42	34.4	395	2	US-09-080-538-3		Sequence 3, Appli
13	42	34.4	420	4	US-09-080-271A-4		Sequence 4, Appli
14	42	34.4	420	4	US-08-974-691-8		Sequence 8, Appli
15	42	34.4	420	4	US-09-705-448-1		Sequence 1, Appli
16	42	34.4	1027	4	US-09-162-021B-2		Sequence 2, Appli
17	41	33.6	17	4	US-08-469-660A-520		Sequence 520, App
18	41	33.6	485	1	US-08-453-356-15		Sequence 15, Appl
19	41	33.6	485	1	US-08-086-631-15		Sequence 15, Appl
20	41	33.6	485	2	US-08-452-930-15		Sequence 15, Appl
21	41	33.6	485	5	PCV-US93-0817A-15		Sequence 15, Appl
22	41	33.6	527	4	US-09-028-366-2		Sequence 2, Appli
23	41	33.6	527	4	US-09-028-366-3		Sequence 3, Appli
24	41	33.6	708	2	US-08-576-165-2		Sequence 2, Appli
25	40	32.8	32	4	US-09-149-476-585		Sequence 585, App
26	40	32.8	372	4	US-09-071-035-102		Sequence 102, App
27	40	32.8	396	1	US-08-176-427B-4		Sequence 4, Appli

RESULT 2
US 08-974-691-6
Sequence 6 Application US/08974691
Sequence 6 6225103
GENERAL INFORMATION:
APPLICANT: Keolsch, Gerald
APPLICANT: Lin, Xindi
APPLICANT: Tang, Jordan
TITLE OF INVENTION: Cloning and Characterization of Napsin
NUMBER OF SEQUENCES: 14

REFERENCE/DOCKET NUMBER: OMRF 166

TELEPHONE: 404-873-8794

TELEFAX: 404-873-8795

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-974-691-3

Query Match 36.9%; Score 45; DB 4; Length 419;

Best Local Similarity 33.3%; Pred. No. 30;

Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 5 QTHCALALGLALATCSOGAA 25

Db 254 QVHMSVKVGTGLSLCAQGS 274

RESULT 5

US-09-705-448-10

Sequence 10, Application US/09705448

Patent No. 6432690

GENERAL INFORMATION:

APPLICANT: Xu, Hong

APPLICANT: Bruno, Sandra A.

APPLICANT: Elsenboss, Laura A.

APPLICANT: Fogliano, Michael

APPLICANT: Cohan, Victoria L.

APPLICANT: Bandman, Olga

TITLE OF INVENTION: HUMAN ASPARTIC PROTEASES

FILE REFERENCE: PF-0458-1 CIP

CURRENT APPLICATION NUMBER: US/09/705,448

CURRENT FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: 09/116,641

PRIOR FILING DATE: 1998-07-16

PRIOR APPLICATION NUMBER: 09/008,271

PRIOR FILING DATE: 1998-01-16

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 10

LENGTH: 419

TYPE: PRT

ORGANISM: MUS MUSCULUS

FEATURE:

OTHER INFORMATION: 1906810, GenBank

US-09-705-448-10

Query Match 36.9%; Score 45; DB 4; Length 419;

Best Local Similarity 33.3%; Pred. No. 30;

Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 5 QTHCALALGLALATCSOGAA 25

Db 254 QVHMSVKVGTGLSLCAQGS 274

RESULT 6

US-08-637-670-25

Sequence 25, Application US/08637670

Patent No. 6413521

GENERAL INFORMATION:

APPLICANT: MCMICHAEL-PHILLIPS et al.

TITLE OF INVENTION: Helminth Parasite Antigen with

SEQUENCE CHARACTERISTICS: Aminopeptidase-like Activity

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Barbara G. Ernst

STREET: 555 13TH STREET, NW Suite 701E

CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/637,670
APPLICATION NUMBER: US/08/637,670
FILING DATE: 26-JUN-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ernst, Barbara G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1811-0232
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-783-6040
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-637-670-25

Query Match 36.1%; Score 44; DB 4; Length 350;
Best Local Similarity 42.1%; Pred. No. 34;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSQIHCALALGLALAI 19

Db 12 LRLTPTLSIALGLIAVAV 30

RESULT 7

US-09-105-343A-2

Sequence 2, Application US/09105343A

Patent No. 6207642

GENERAL INFORMATION:

APPLICANT: WILEY, S.R.

TITLE OF INVENTION: MEMBER OF THE TNF FAMILY USEFUL

FOR TREATMENT AND DIAGNOSIS OF DISEASE

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-6050

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,343A

FILING DATE: 12-FEB-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/02859

FILING DATE: 12-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: BECKER, Cheryl E.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6048.US.P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-935-1729

TELEFAX: 847-938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6207642e
US-09-105-343A-2

Query Match 35.2%; Score 43; DB 4; Length 249;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Caps 0;

QY 9 ALALLGLALAICSOGA 24
Db 31 ALACGLGLLVLSGS 46

RESULT 8
US-08-469-667-14
Sequence 14, Application US/08469667
Patent No. 5733748
GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-667-14

Query Match 34.4%; Score 42; DB 1; Length 81;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Caps 0;

QY 9 ALALLGLALAICSOGA 25
Db 5 ALCMGLGLVLLSSSSA 21

RESULT 9
US-09-224-110-14
Sequence 14, Application US/09224110
Patent No. 6337195
GENERAL INFORMATION:

APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-435
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-224-110-14

Query Match 34.4%; Score 42; DB 4; Length 81;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Caps 0;

QY 9 ALALLGLALAICSOGA 25
Db 5 ALCMGLGLVLLSSSSA 21

RESULT 10
PCT-US95-07289-14
Sequence 14, Application PC/TUS9507289
GENERAL INFORMATION:
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Colon Specific Genes and Proteins
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: NJ
COUNTRY: USA
ZIP: 07068-1739

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 06-JUN-1995
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-365
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT:US95-07289-14

Query Match 34.4%; Score 42; DB 5;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 5; Indels

QY 9 ALALLGLALAICSGAA 25
|| : ||| || : |
Db 5 ALCMLGLVLLALLSSSA 21

```

RESULT 11
US-08-723-938-3
; Sequence 3, Application US/08723938
; Patent No. 5776759
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS
; NUMBER OF SEQUENCES: 4

```

TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
IMMEDIATE SOURCE:
LIBRARY: LUNGNOT02

CLONE: 312099
US-08-723-938-3

Query Match 34.4%; Score 42; DB 1; Length 395;
Best Local Similarity 38.1%; Pred. No. 80;
Matches 8; Conservative 4; Mismatches 9; Indels

QY 5 QIHCALALLGLALAICSQGAA 25
 ||| : | | : : | |
 Db 259 QIHMERVKVGPGGLTLCAGKCA 279

RESULT 12
US-09-080-538-3
; Sequence 3, Application US/09080538
; Patent No. 5965129
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS
; NUMBER OF SEQUENCES: 4

Query Match	34.4%	Score 42;	DB 2;	Length 395;
Best Local Similarity	38.1%	Pred. No. 80;		
Matches	8;	Conservative	4;	Mismatches
				9; Indels

QY 5 QIHCALALLGLALAICSOGAA 25
 ||| :| | :| :| |
 Db 259 OIHMERYKVGPGITLCAKGA 279

RESULT 13

US-09-008-271A-4
Sequence 4, Application US/09008271A
Patent No. 6203979
GENERAL INFORMATION:

APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
Shah, Tom Y.
Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/008,271A

FILING DATE: 16-Jan-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mohan-Peterson, Sheila

REGISTRATION NUMBER: 41,201

REFERENCE/DOCKET NUMBER: PF-0458 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-835-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 420 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGAST01

CLONE: 877617

SEQUENCE DESCRIPTION: SEQ ID NO: 4 :

US-09-008-271A-4

Query Match 34.4%; Score 42; DB 4; Length 420;
Best Local Similarity 38.1%; Pred. No. 85;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 QIHCALALGLALATCSOGAA 25

|||||

Db 259 QIHMERVKVPGTLCAKCA 279

RESULT 14

US-08-974-691-8
Sequence 8, Application US/08974691
Patent No. 6225103
GENERAL INFORMATION:

APPLICANT: Keolsch, Gerald

APPLICANT: Lin, Xinli

APPLICANT: Tang, Jordan

TITLE OF INVENTION: Cloning and Characterization of Napsin

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center, 1201 W. Peachtree

STREET: St.
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/974,691
APPLICATION NUMBER: US/08/974,691
FILING DATE: 20-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA: US 60/031,196
APPLICATION NUMBER: US 60/031,196
FILING DATE: 20-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,126
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF 166
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-873-8794
TELEFAX: 404-873-8795
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-974-691-8

Query Match 34.4%; Score 42; DB 4; Length 420;
Best Local Similarity 38.1%; Pred. No. 85;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 QIHCALALGLALATCSOGAA 25

|||||

Db 259 QIHMERVKVPGTLCAKCA 279

RESULT 15

US-09-705-448-1
Sequence 1, Application US/09705448
Patent No. 6432690
GENERAL INFORMATION:

APPLICANT: Xu, Hong

APPLICANT: Bruno, Sandra A.

APPLICANT: Elsenboss, Laura A.

APPLICANT: Fogliano, Michael

APPLICANT: Cohen, Victoria L.

APPLICANT: Bardman, Olga

TITLE OF INVENTION: HUMAN ASPARTIC PROTEASES

FILE REFERENCE: PF-0458-1 CIP

CURRENT APPLICATION NUMBER: US/09/705,448

CURRENT FILING DATE: 2000-11-02

PRIOR FILING DATE: 1998-07-16

PRIOR APPLICATION NUMBER: 09/116,641

PRIOR FILING DATE: 1998-07-16

PRIOR APPLICATION NUMBER: 09/008,271

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 420

TYPE: PRT

ORGANISM: HOMO SAPIENS

FEATURE:

OTHER INFORMATION: 372637, LUNGNOT02
US-09-705-448-1

Query Match 34.4%; Score 42; DB 4; Length 420;
Best Local Similarity 38.1%; Pred. No. 85;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 5 QIHCALALLGLALAICSGAA 25
||| .| | | | | | | | | |
Db 259 QIHMERVKVGPGLTLCAGCA 279

Search completed: March 21, 2003, 11:42:36
Job time : 5.72103 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 21, 2003, 14:35:23 ; Search time 15.4506 seconds
(without alignments)
215.607 Million cell updates/sec

Title: US-09-727-739b-19

Perfect score: 122

Sequence: 1 MRVSIHCALALLGLAICSGAA 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues.

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```
1: A_Geneseq_101002.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	111	22	AAU07668
2	110	90.2	115	22	AAU07667
3	64	52.5	114	22	AAU07666
4	52	42.6	433	18	AAW14001
5	52	42.6	434	23	AAW48922
6	48	39.3	120	22	AAW81913
7	48	39.3	162	22	AAW81978
8	48	39.3	429	21	AAW42064
9	48	39.3	433	19	AAW54357
10	48	39.3	434	23	ABB57379

11	47	38.5	210	20	AAV35223	Chlamydia pneumoni
12	47	38.5	423	23	ABA47765	Novel G-protein co
13	47	38.5	472	23	AAU93172	Arbidopsis transc
14	47	38.5	475	23	AAU93018	Arbidopsis transc
15	47	38.5	481	23	ABA47764	Novel G-protein co
16	47	38.5	502	23	ABA47767	Novel G-protein co
17	47	38.5	560	23	ABG34861	Human G-protein co
18	47	38.5	560	23	ABA47766	Novel G-protein co
19	47	38.5	574	22	ABA94618	Human protein sequ
20	47	38.5	682	23	AAO15109	Human G protein-co
21	47	38.5	718	22	AAO92973	Human protein sequ
22	47	38.5	724	22	ABG17176	Novel human diagn
23	47	38.5	769	23	ABA47761	Novel G-protein co
24	47	38.5	797	22	AAO95679	Human protein sequ
25	47	38.5	797	23	AAE23412	Human G-protein co
26	47	38.5	798	22	AAO93238	Human protein sequ
27	47	38.5	848	23	ABA47763	Novel G-protein co
28	47	38.5	1150	22	AAW40294	Human polypeptide
29	46	37.7	400	22	AAW72561	Rat G protein-coup
30	46	37.7	400	23	AAO14445	Rattus norvegicus
31	46	37.7	433	21	AAW44810	Human Aspartic pro
32	46	37.7	438	19	AAW57042	Human aspartic pro
33	46	37.7	451	19	AAW54877	Human napsin A pro
34	46	37.7	1982	22	ABG09731	Novel human diagn
35	45	36.9	311	21	RAY44937	Corn sulphate perm
36	45	36.9	395	21	ANG14553	Arbidopsis thalia
37	45	36.9	395	21	ANG52413	Arbidopsis thalia
38	45	36.9	419	21	RAY79177	Haemacopoietic ste
39	45	36.9	425	21	ANG14552	Arbidopsis thalia
40	45	36.9	425	21	ANG52412	Arbidopsis thalia
41	45	36.9	453	21	ANG44551	Arbidopsis thalia
42	45	36.9	453	21	ANG52411	Arbidopsis thalia
43	45	36.9	453	23	AAO93186	Arbidopsis transc
44	45	36.9	1039	22	ABB56566	Drosophila melanog
45	44.5	36.5	116	21	AAG03774	Human secreted pro

ALIGNMENTS

RESULT 1
AAU07668
ID AAU07668 standard; Protein; 111 AA.

XX AAU07668:

XX 04-DEC-2001 (first entry)

XX Rainbow trout preprosomatostatin II (PPSS-II'') polypeptide.

XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
PPSS-II'; PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
KW carcinoma syndrome; cell proliferation; apoptosis; growth hormone;
KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
KW neoplasia; Alzheimer's disease; Huntington's disease; neuroprotective;
KW epilepsy; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
XX anti-human immunodeficiency virus; osteopathic; anticonvulsant.

OS Oncorhynchus mykiss.

Key	Location/Qualifiers
XX	1..25
FT	/note= "signal peptide"
FT	Protein
FT	1..86
FT	/note= "PPSS-II'' pre-sequence"
FT	Protein
FT	26..111
FT	/note= "Mature PPSS-II''"
FT	Peptide
FT	87..97
FT	/note= "PPSS-II'' pro-sequence"
FT	Peptide
FT	87..111
FT	/note= "Prosomatostatin II''"
FT	Cleavage-site
FT	95..97

FT Peptide /note= "Dibasic cleavage site"
 FT 98..111
 FT /note= "SS-14 variant peptide"
 XX CA2325169-A1.
 XX 03-JUN-2001.
 XX 01-DEC-2000; 2000CA-2325169.
 XX 03-DEC-1999; 99US-0168934.
 XX (NDSU-) NDSU RES FOUND.
 XX Sheridan MA, Moore CA, Kittelson JD;
 XX WPI; 2001-425997/46..
 XX N-PSDB; AAS12935.
 XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
 XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired
 XX immunodeficiency syndrome and neurological disorders -
 XX Claim 1; Fig 3; 52pp; English.
 XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
 XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
 XX preprosomatostatin II (PPSS-II). The protein sequences and their
 XX associated polynucleotides are useful for identifying modified
 XX somatostatin polypeptide which functions as a somatostatin agonist useful
 XX for research, therapeutics or diagnostics, including medical and
 XX veterinary applications. The wild-type somatostatin and its modified
 XX version are useful for treating hypersecretion from endocrine tumours in
 XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
 XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
 XX through their effects on cell proliferation and apoptosis and as adjuncts
 XX in the treatment of diabetes mellitus via inhibition of growth hormone
 XX and glucagon. In addition, dysfunctional somatostatin secretion is
 XX associated with acquired immunodeficiency syndrome (AIDS) and various
 XX neurological disorders (e.g. epilepsy, Alzheimer's disease and
 XX Huntington's disease) and somatostatin antagonists are effective in the
 XX treatment of such conditions. Nucleic acids encoding the polypeptides are
 XX useful in gene therapy and fusion peptides can be targeted to neoplasms
 XX and their metastases, inhibiting the release of their secretory products.
 XX This sequence represents O. Mykiss PPSS-II protein.
 XX Note: The features for this sequence are specifically claimed in the
 XX specification.
 XX Sequence 111 AA:
 XX Query Match 100.0%; Score 122; DB 22; Length 111;
 XX Best Local Similarity 100.0%; Pred. No. 1e-10; Mismatches 0; Indels 0; Gaps 0;
 XX Matches 25; Conservative 0;
 OY 1 MRVSIHCAALLGLAIAICSGAA 25
 Db 1 MRVSIHCAALLGLAIAICSGAA 25
 RESULT 2
 AAU07667
 ID AAU07667 standard; Protein; 115 AA.
 XX AAU07667;
 AC AAU07667;
 XX 04-DEC-2001 (first entry)
 DT Rainbow trout preprosomatostatin II (PPSS-II') polypeptide.
 DE Rainbow trout preprosomatostatin; hypersecretion; PPSS-I;
 XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
 KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
 KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
 KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
 KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
 KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
 KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
 KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.
 XX Oncorhynchus mykiss.
 XX Key Location/Qualifiers
 XX Peptide 1..25 /note= "signal peptide"
 XX Protein 1..87 /note= "PPSS-II' pre-sequence"
 XX Protein 26..115 /note= "Mature PPSS-II' "
 XX Misc-difference 74 /note= "Encoded by CAA"
 XX Peptide 88..101 /note= "PPSS-II' pro-sequence"
 XX Peptide 88..115 /note= "Preprosomatostatin II' "
 XX Cleavage-site 100..101 /note= "Dibasic cleavage site"
 XX Peptide 102..115 /note= "SS-14 variant peptide"
 XX CA2325169-A1.
 XX 03-JUN-2001.
 XX 01-DEC-2000; 2000CA-2325169.
 XX 03-DEC-1999; 99US-0168934.
 XX (NDSU-) NDSU RES FOUND.
 XX Sheridan MA, Moore CA, Kittelson JD;
 XX WPI; 2001-425997/46..
 XX N-PSDB; AAS12934.
 XX New somatostatin polypeptides derived from Oncorhynchus mykiss, useful
 XX for treating diabetes mellitus, acromegaly, gastrinoma, acquired
 XX immunodeficiency syndrome and neurological disorders -
 XX Claim 2; Fig 3; 52pp; English.
 XX The invention relates to an Oncorhynchus mykiss somatostatin polypeptide
 XX containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
 XX preprosomatostatin II (PPSS-II). The protein sequences and their
 XX associated polynucleotides are useful for identifying modified
 XX somatostatin polypeptide which functions as a somatostatin agonist useful
 XX for research, therapeutics or diagnostics, including medical and
 XX veterinary applications. The wild-type somatostatin and its modified
 XX version are useful for treating hypersecretion from endocrine tumours in
 XX the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
 XX gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
 XX through their effects on cell proliferation and apoptosis and as adjuncts
 XX in the treatment of diabetes mellitus via inhibition of growth hormone
 XX and glucagon. In addition, dysfunctional somatostatin secretion is
 XX associated with acquired immunodeficiency syndrome (AIDS) and various
 XX neurological disorders (e.g. epilepsy, Alzheimer's disease and
 XX Huntington's disease) and somatostatin antagonists are effective in the
 XX treatment of such conditions. Nucleic acids encoding the polypeptides are
 XX useful in gene therapy and fusion peptides can be targeted to neoplasms
 XX and their metastases, inhibiting the release of their secretory products.
 XX This sequence represents O. Mykiss PPSS-II protein.
 XX Note: The features for this sequence are specifically claimed in the
 XX specification.
 XX Sequence 115 AA:
 XX Query Match 90.2%; Score 110; DB 22; Length 115;
 XX Best Local Similarity 88.0%; Pred. No. 6.7e-09;

Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRVSOIHCALALGLALATCSOGAA 25
 DB 1 MKVCRIRHCALALGLALATCSOGAA 25

RESULT 3
 AAU07666
 ID AAU07666 standard; Protein; 114 AA.
 AC AAU07666;
 DT 04-DEC-2001 (first entry)
 XX Rainbow trout preprosomatostatin I (PPSS-I) polypeptide.

XX Rainbow trout; somatostatin; preprosomatostatin; hypersecretion; PPSS-I;
 KW PPSS-II'; endocrine tumour; pituitary gland; glucagonoma; AIDS;
 KW gastroenteropancreatic tissue; acromegaly; gastrinoma; diabetes mellitus;
 KW carcinoid syndrome; cell proliferation; apoptosis; growth hormone;
 KW glucagon; acquired immunodeficiency syndrome; neurological disorder; HIV;
 KW epilepsy; Alzheimer's disease; Huntington's disease; neuroprotective;
 KW neoplasm; metastasis; gene therapy; antidiabetic; nootropic; cytostatic;
 KW anti-human immunodeficiency virus; osteopathic; anticonvulsant.

XX Oncochrychus mykiss.

XX Key Location/Qualifiers
 FH Peptide 1..24
 FT /note= "Signal peptide"
 FT Protein 1..88
 FT /note= "PPSS-I pre-sequence"
 FT Protein 25..114
 FT /note= "Mature PPSS-I"
 FT Peptide 89..100
 FT /note= "PPSS-I pro-sequence"
 FT Peptide 89..114
 FT /note= "Prosomatostatin I"
 FT Cleavage-site 99..100
 FT /note= "Dibasic cleavage site"
 FT Peptide 101..114
 FT /note= "SS-14 peptide"

XX CA2325169-A1.
 PN 03-JUN-2001.
 PD 01-DEC-2000; 2000CA-2325169.
 PF 03-DEC-1999; 99US-0168934.
 PR (NDSU-) NDSU RES FOUND.
 XX Sheridan MA, Moore CA, Kittelson JD;
 PI WPI: 2001-425997/46.
 XX N-PSDB; AAS12933.

XX New somatostatin polypeptides derived from Oncochrychus mykiss, useful
 for treating diabetes mellitus, acromegaly, gastrinoma, acquired
 immunodeficiency syndrome and neurological disorders.

XX Claim 1; Fig 2; 52pp; English.

XX The invention relates to an Oncochrychus mykiss somatostatin polypeptide
 containing a portion of preprosomatostatin I (PPSS-I) and/or a portion of
 preprosomatostatin II (PPSS-II). The protein sequences and their
 associated polynucleotides are useful for identifying modified
 somatostatin polypeptides which functions as a somatostatin agonist
 for research, therapeutics or diagnostics, including medical and
 veterinary applications. The wild-type somatostatin and its modified
 version are useful for treating hypersecretion from endocrine tumours in

CC the pituitary (e.g. acromegaly) or gastroenteropancreatic tissues (e.g.
 CC gastrinoma, glucagonoma, carcinoid syndrome), to cause tumour shrinkage
 CC through their effects on cell proliferation and apoptosis and as adjuncts
 CC in the treatment of diabetes mellitus via inhibition of growth hormone
 CC and glucagon. In addition, dysfunctional somatostatin secretion is
 CC associated with acquired immunodeficiency syndrome (AIDS) and various
 CC neurological disorders (e.g. epilepsy, Alzheimer's disease and
 CC Huntington's disease) and somatostatin antagonists are effective in the
 CC treatment of such conditions. Nucleic acids encoding the polypeptides are
 CC useful in gene therapy and fusion peptides can be targeted to neoplasms
 CC and their metastases, inhibiting the release of their secretory products.
 CC This sequence represents O. Mykiss PPSS-I protein.
 CC Note: The features for this sequence are specifically claimed in the
 CC specification.

XX Sequence 114 AA;

Query Match 52.5%; Score 64; DB 22; Length 114;
 Best Local Similarity 60.0%; Pred. No. 0.048; 7; Indels 0; Gaps 0;
 Matches 13; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRVSOIHCALALGLALATCSOGAA 25
 DB 1 MLSTRVOCALALLSLALISSVSA 25

RESULT 4
 AAW14001
 ID AAW14001 standard; Protein; 433 AA.
 XX AC AAW14001;
 XX DT 23-MAY-1997 (first entry)
 XX DE Enolase protein.
 XX KW PCR: polymerase chain reaction; primer: amplify; tyrosine; human; NSE;
 KW enolase; radioisotope; antigen.
 XX OS Homo sapiens.
 XX JF08308584-A.
 XX 26-NOV-1996.
 XX PF 19-MAY-1995; 95JP-0145542.
 XX PR 19-MAY-1995; 95JP-0145542.
 XX PA (EIKE) EIKEN KAGAKU KK.
 XX WPI: 1997-059703/06.
 XX Introducing tyrosine residues into a protein for radioisotopic
 PT labelling - by substitution, addition and/or insertion to a DNA
 PT coding for the protein, antigenicity of the protein is unchanged by
 PT labelling
 XX PS Disclosure; Page 10-11; 13pp; Japanese.
 XX This sequence represents the wild type human enolase (NSE) protein. The
 CC primers represented by AAT60282-T60285 were used to amplify tyrosine
 CC introduced versions of this sequence. The introduced tyrosine residues
 CC combine a radioisotope in such a way as to have no substantial effect on
 CC the antigenicity of the protein. The peptides with the introduced
 CC tyrosine residue are prepared by expressing a recombinant DNA sequence,
 CC where the tyrosine residue has been introduced via an insertion or
 CC substitution into the wild type sequence. The peptides can then be used
 CC in radioimmunoassays, as the introduction of the label does not alter
 CC the antigenicity of the protein.

XX Sequence 433 AA;

Query Match 42.6%; Score 52; DB 18; Length 433;
 Best Local Similarity 58.8%; Pred. No. 12;
 Matches 10; Conservative 4; Mismatches 0; Indels 3; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
 | : : : : : : : : : :
 Db 107 ANAILGVSLAVCKAGAA 123

RESULT 5

AAAM48922
 ID AAAM48922 standard; Protein; 434 AA

AC AAAM48922;

DT 16-APR-2002 (first entry)

DE Human enolase 3(beta,muscle) protein.

KW Human; enolase 3(beta, muscle); ENO3; single nucleotide polymorphism;
 KW SNP; haplotype analysis; isogene.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 71 /note= "optionally Ser"
 FT Misc-difference 76 /note= "optionally Pro"
 FT Misc-difference 85 /note= "optionally Val"
 FT Misc-difference 332 /note= "optionally Ile"

XX WO200202579-A2.

PN 10-JAN-2002.

XX 02-JUL-2001; 2001WO-US20952.

XX 30-JUN-2000; 2000US-215236P.

XX (GENA-) GENAISANCE PHARM INC.

XX Duda A, Finkel X, Koshy B, Parks KE;

XX WPI; 2002-154721/20.

XX N-PSDB; AAK98531, AAK98532.

XX Novel genetic variants of enolase 3, (beta, muscle) gene useful in
 XX studying expression and function of the protein, and for screening
 XX drugs to treat disorders of glycolytic pathway

PS Claim 28; Fig 3; 90pp; English.

XX The present invention provides the protein, cDNA and genomic sequences of
 CC a human enolase 3 (beta, muscle) isogene containing a number of single
 CC nucleotide polymorphisms (SNPs). The sequences can be used to identify
 CC the haplotype of an individual and identify whether particular haplotypes
 CC are linked to certain diseases. The present sequence is the ENO3
 CC protein.

XX Sequence 434 AA;

Query Match 42.6%; Score 52; DB 23; Length 434;
 Best Local Similarity 58.8%; Pred. No. 12;
 Matches 10; Conservative 4; Mismatches 0; Indels 3; Gaps 0;

QY 9 ALALLGLALAICSGAA 25
 | : : : : : : : : : :
 Db 108 ANAILGVSLAVCKAGAA 124

RESULT 6

AAAM81913
 ID AAAM81913 standard; Protein; 120 AA

XX AC AAAM81913;

XX 13-NOV-2001 (first entry)

XX Human haematological malignancy-related antigen #1611.

XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
 KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
 KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.

OS Homo sapiens.

XX WO200164886-A2.

XX 07-SEP-2001.

XX 01-MAR-2001; 2001WO-US07272.

XX 01-MAR-2000; 2000US-0186126.

XX 17-MAR-2000; 2000US-0190479.

XX 27-APR-2000; 2000US-0200545.

XX 28-APR-2000; 2000US-0200303.

XX 28-APR-2000; 2000US-0200779.

XX 01-MAY-2000; 2000US-0200999.

XX 04-MAY-2000; 2000US-0202084.

XX 14-JUL-2000; 2000US-0206201.

XX 03-AUG-2000; 2000US-0218950.

XX 04-AUG-2000; 2000US-0222903.

XX 07-AUG-2000; 2000US-0223416.

XX (CORI-) CORIXA CORP.

XX Gaiger A, Algate PA, Mannion J;

XX WPI; 2001-514842/56.

XX Compositions and methods for the detection of hematological
 XX malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
 XX lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma

XX Claim 1; Page 1152; 1252pp; English.

XX The present invention relates to compositions and methods for the
 CC detection, diagnosis and therapy of haematological malignancies. The
 CC present sequence is the protein sequence of a human haematological
 CC malignancy related antigen. The methods of the present invention comprise
 CC detecting the presence of haematological malignancy related antigen(s) in
 CC a sample obtained from the patient (an increased level of the
 CC polypeptide, compared to an unaffected individual, is indicative of an
 CC increased risk). Haematological malignancies which can be treated using
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
 CC cell non-Hodgkin's lymphoma.

XX Sequence 120 AA;

Query Match 39.3%; Score 48; DB 22; Length 120;
 Best Local Similarity 56.2%; Pred. No. 12;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 ALAILGLALAICSGAA 24

Db 55 ANAILGVSLAVCKAGA 80

RESULT 7

AAAM81978
 ID AAAM81978 standard; Protein; 162 AA


```

OY 9 ALALGLALATCSOGA 24
DB 103 ANAILGVSLAVCKAGA 118

RESULT 9
AAW54357
ID AAW54357 standard; protein; 433 AA.
AC
XX
XX
XX 14-AUG-1998 (first entry)
XX Alpha Enolase.
XX Endometrium; hyperplasia; adenocarcinoma; proliferative phase;
KW 2D gel electrophoresis; detection.
OS Homo sapiens.
XX
XX WO9810291-A1.
XX
XX 12-MAR-1998.
XX
XX 05-SEP-1997; 97WO-G802394.
XX
XX 08-APR-1997; 97GB-0007132.
XX 06-SEP-1996; 96GB-0018600.
XX
XX (CLIN-) CENT CLINICAL & BASIC RES.
XX
XX Byrjalsen I, Fey SJ, Larsen P;
XX WPI; 1998-207057/18.
XX
XX Biochemical markers of human endometrium - useful for, e.g.
XX diagnosis of hyperplasia and adenocarcinoma
XX
XX Disclosure; Page 21; 77pp; English.
XX
XX Proteins AAW54349-H54364 are examples of proteins produced in the
XX endometrium during the hyperplasia, adenocarcinoma or proliferative
XX phase of the endometrium. The presence and quantities of these proteins
XX can be detected using 2D gel electrophoresis comparison of cell lysates.
XX The proteins can be used as biochemical markers to detect the phase of
XX the endometrium and can be measured in body fluids, obviating the need
XX for endometrial biopsies.
XX
XX SQ Sequence 433 AA;
Query Match 39.3%; Score 48; DB 19; Length 433;
Best Local Similarity 56.2%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 9 ALALGLALATCSOGA 24
DB 107 ANAILGVSLAVCKAGA 122

RESULT 10
ABB57379
ID ABB57379 standard; protein; 434 AA.
AC
XX
XX
XX 08-MAR-2002 (first entry)
XX
XX Rat mucocardial cell proliferation associated polypeptide SEQ ID NO 10.
XX
XX Rat; heart; cardiant; myocardial necrosis; cardiac hypertrophy;
XX cardiac insufficiency.
XX
XX Rattus norvegicus.
XX
XX PI

OY 9 ALALGLALATCSOGA 24
DB 108 ANAILGVSLAVCKAGA 123

RESULT 11
AAV35223
ID AAV35223 standard; protein; 210 AA.
XX
XX
XX
XX 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae protein not found in C. trachomatis.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
XX
XX WO9927105-A2.
XX
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX 21-NOV-1997; 97FR-0014673.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
XX PI

OY 9 ALALGLALATCSOGA 24
DB 108 ANAILGVSLAVCKAGA 123

Query Match 39.3%; Score 48; DB 23; Length 434;
Best Local Similarity 56.2%; Pred. No. 49;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 9 ALALGLALATCSOGA 24
DB 108 ANAILGVSLAVCKAGA 123

RESULT 11
AAV35223
ID AAV35223 standard; protein; 210 AA.
XX
XX
XX
XX 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae protein not found in C. trachomatis.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
XX
XX WO9927105-A2.
XX
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
XX 21-NOV-1997; 97FR-0014673.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
XX PI

```

XX	WPI; 1999-357842/30.
DR	Genome sequence of Chlamydia pneumoniae.
XX	Page 1070-1071; Disclosure; 1912pp; English.
XX	AAV34584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC	C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see AAV34584-Y35879) can be used in immunogenic compositions as vaccines. Vectors containing C. pneumoniae nucleotide sequences can also be used as immunogenic compositions, especially where the vector directs the expression of a neutralising epitope of C. pneumoniae.
CC	Sequence 210 AA:
XX	Query Match 38.5%; Score 47; DB 20; Length 210;
XX	Best Local Similarity 45.0%; Pred. No. 32;
XX	Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps
QY	5 QIHCALALLGLALAICSGA 24
DBD	48 EIASAIALGLLVAFCSAA 67
RESULT 12	
ABE47765	
ID AAB47765 standard; Protein; 423 AA.	
AC AAB47765;	
DT DT	21-FEB-2002 (first entry)
DE XX	Novel G-protein coupled receptor #5.
KK KW	Novel G-protein coupled receptor; NGPCR; transmembrane domain;
KW KK	signal transduction; ligand binding; human; placenta; bone marrow;
KW KK	trachea; testis; liver; kidney cell.
OS	Homo sapiens.
FN PN	WO200181411-A2.
PD PD	01-NOV-2001.
XX XX	27-APR-2001; 2001WO-US13371.
XX XX	27-APR-2000; 2000US-199950P.
PP PR	(LEXI-) LEXICON GENETICS INC.
PA PA	Wang X, Walke DW, Scoville J;
PI PI	WPI: 2002-049266/06.
DR DR	N-PSDB; AA172010.
PT PT	Novel isolated G-protein coupled receptor polypeptide useful for identifying compounds that modulate the protein expression or activity which are used to treat a variety of physiological or mental disorders
PS PS	Disclosure; Page 78-79; 85pp; English.
XX CC	The sequences given in AAB47761-69 show novel G-protein coupled receptor (NGPCR) polypeptides. The NGPCR coding sequences are found on chromosome 3 at 3q11.1 in the human genome. These NGPCR proteins show the presence of 7 conserved transmembrane domains which are interconnected by non-conserved hydrophilic loops. They are involved

CC In signal transduction after ligand binding. These NPCR's are
CC expressed in human placenta, bone marrow, trachea, testis, liver and
CC kidney cells.
XX
SQ Sequence 423 AA;

Query Match 38.5%; Score 47; DB 23; Length 423;
Best Local Similarity 45.0%; Pred. No. 67;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0

OY 3 VSQIHCALALLGLALAICSQ.22
DB 142 LSNVGCALSVTGLALTIVFQ 161
| : | : | : | : | : | :
| : | : | : | : | : | :

RESULT 13
AAU93172
ID AAU93172 standard; Protein: 472 AA.
XX AC
XX AAU93172:
DT 02-JUL-2002 (first entry)
DE XX Arabidopsis transcription factor #210.
KW Agriculture; metabolic chemical; environmental stress; drought;
KW microbial disease resistance; herbicide resistance; seed yield;
KW fruit yield; growth rate; leaf senescence; flower senescence.
KW Plant; transcription factor; transgenic.
XX OS Arabidopsis thaliana.
XX WO200215675-A1.
XX PN
XX PD 28-FEB-2002.
XX PF 22-AUG-2001; 2001WO-US26189.
XX PP 22-AUG-2000; 2000US-227439P.
PR 16-NOV-2000; 2000US-071399A.
PR 16-APR-2001; 2001US-083794A.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (PILG/) PILGRIM M.
PA (CREE/) CREELMAN R.
PA (DUBE/) DUBELL A J.
PA (HEAR/) HEARD J.
PA (JIAN/) JIANG C.
PA (KEDD/) KEDDIE J.
PA (ADAM/) ADAM L.
PA (RATC/) RATCLIFF O.
PA (REUB/) REUBER J L.
PA (RIEC/) RIECHMANN J L.
PA (YUGC/) YU G.
PA (PINE/) PINEDA O.

Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,
Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
WPI: 2002-292022/33.
DR N-PSDB; ABK65358.

An isolated or recombinant polynucleotide used to produce a transgenic plant -
XX
PPS Claim 40; Page 864-866; 94lpp; English.

The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant

CC exhibits ectopic expression or altered expression of one or more genes
 CC associated with a plant trait as compared to a wild plant. Also included
 CC are a transgenic plant comprising the polynucleotides, a computer
 CC readable medium having stored sequence information, and identifying a
 CC homologue sequence from a database comprising a plurality of known plant
 CC sequences comprising inputting sequence information selected from one of
 CC 464 fully defined sequences given in the specification. The isolated or
 CC recombinant polynucleotide is used for producing a plant having a
 CC modified trait. The method comprising selecting a polynucleotide that
 CC encodes a polypeptide or an antisense nucleic acid, inserting the
 CC polynucleotide or antisense nucleic acid into an expression vector,
 CC introducing the vector into a plant or a cell of a plant to overexpress
 CC the polypeptide or antisense nucleic acid, thereby producing a modified
 CC plant, and selecting for modified trait (e.g. increased chemical,
 CC production of agriculturally useful proteins or metabolic chemicals,
 CC pest tolerance, environmental stress response (e.g. drought), microbial
 CC disease resistance, herbicide resistance, seed and fruit yield, growth
 CC rate, leaf and flower senescence and many other traits listed in the
 CC specification). The present sequence is one of the 232 proteins which are
 CC A. thaliana transcription factors.

XX Sequence 472 AA;
 Query Match 38.5%; Score 47; DB 23; Length 472;
 Best Local Similarity 40.9%; Pred. No. 75;
 Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SQIHCALALLGLALALCSOGAA 25
 Db 193 AQIHAASVAGVAAVAATAAA 214
 :||| |::: |::| : ||

RESULT 14
 AAU93018
 ID AAU93018 standard; Protein: 475 AA.
 AC AAU93018;
 XX

DT 02-JUL-2002 (first entry)

DE Arabidopsis transcription factor #56.

KW Agriculture; metabolic chemical; environmental stress; drought;
 KW microbial disease resistance; herbicide resistance; seed yield;
 KW fruit yield; growth rate; leaf senescence; flower senescence.
 KW plant; transcription factor; transgenic.

OS Arabidopsis thaliana.

PN WO200215675-A1.

XX 28-FEB-2002.

XX 22-AUG-2001; 2001WO-US26189.

PR 22-AUG-2000; 2000US-227439P.

PR 16-NOV-2000; 2000US-0713994.

PR 16-APR-2001; 2001US-0637944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PILG/) PILGRIM M.

PA (CREE/) CREELMAN R.

PA (DUBE/) DUBELL A J.

PA (HEAR/) HEARD J.

PA (JIAN/) JIANG C.

PA (KEDD/) KEDDIE J.

PA (ADAM/) ADAM L.

PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.

PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
 XX WPI: 2002-292022/33.

DR N-PSDB; ABK65204.

XX An isolated or recombinant polynucleotide used to produce a transgenic

PT plant

XX Claim 40; Page 265-267; 941pp; English.

XX The invention relates to 1 of 232 isolated or recombinant polynucleotides

CC encoding an Arabidopsis thaliana transcription factor, that variants,

CC complements, fragments, or related polynucleotide with 31% as compared

CC to a wild-type or reference plant, or the plant possesses an altered

CC phenotype as compared to a wild-type or reference plant, or the plant

CC exhibits ectopic expression or altered expression of one or more genes

CC associated with a plant trait as compared to a wild plant. Also included

CC are a transgenic plant comprising the polynucleotides, a computer

CC readable medium having stored sequence information, and identifying a

CC homologue sequence from a database comprising a plurality of known plant

CC sequences comprising inputting sequence information selected from one of

CC 464 fully defined sequences given in the specification. The isolated or

CC recombinant polynucleotide is used for producing a plant having a

CC modified trait. The method comprising selecting a polynucleotide that

CC encodes a polypeptide or an antisense nucleic acid, inserting the

CC polynucleotide or antisense nucleic acid into an expression vector,

CC introducing the vector into a plant or a cell of a plant to overexpress

CC the polypeptide or antisense nucleic acid, thereby producing a modified

CC plant, and selecting for a modified trait (e.g. increased chemical,

CC production of agriculturally useful proteins or metabolic chemicals,

CC pest tolerance, environmental stress response (e.g. drought), microbial

CC disease resistance, herbicide resistance, seed and fruit yield, growth

CC rate, leaf and flower senescence and many other traits listed in the

CC specification). The present sequence is one of the 232 proteins which are

CC A. thaliana transcription factors.

XX Sequence 475 AA;

Query Match 38.5%; Score 47; DB 23; Length 475;

Best Local Similarity 40.9%; Pred. No. 75;

Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SQIHCALALLGLALALCSOGAA 25

Db 198 AQIHAASVAGVAAVAATAAA 219

:||| |::: |::| : ||

RESULT 15

AAU93018

ID AAB47764 standard; Protein: 481 AA.

XX AAB47764;

XX AAB47764;

XX 21-FEB-2002 (first entry)

XX Novel G-protein coupled receptor #4.

XX Novel G-protein coupled receptor; NGPCR; transmembrane domain;

XX signal transduction; ligand binding; human; placenta; bone marrow;

XX trachea; testis; liver; kidney cell.

XX Homo sapiens.

XX WO200181411-A2.

XX 01-NOV-2001.

XX 27-APR-2001; 2001WO-US13371.

XX 27-APR-2000; 2000US-199950P.

XX

PA (LEXI-) LEXICON GENETICS INC.
 XX
 FI Wang X, Walke DW, Scoville J;
 XX
 DR WPI: 2002-049266/06.
 DR N-PSDB; AA172009.
 XX
 PT Novel isolated G-protein coupled receptor polypeptide useful for
 PT identifying compounds that modulate the protein expression or activity
 PT which are used to treat a variety of physiological or mental disorders
 PT
 XX
 PS Claim 6; Page 77; 85pp; English.
 XX
 CC The sequences given in AAB47761-69 show novel G-protein coupled
 CC receptor (NGPCR) polypeptides. The NGPCR coding sequences are found on
 CC chromosome 3 at 3q11.1 in the human genome. These NGPCR proteins show
 C the presence of 7 conserved transmembrane domains which are
 CC interconnected by non-conserved hydrophilic loops. They are involved
 CC in signal transduction after ligand binding. These NGPCR's are
 CC expressed in human placenta, bone marrow, trachea, testis, liver and
 CC kidney cells.
 XX
 SQ Sequence 481 AA:

Query Match 38.5%; Score 47; DB 23; Length 481;
 Best Local Similarity 45.0%; Pred. No. 76;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 VSOIHCALALLGLAICSQ 22
 Db :|::||::||::|
 Db 200 LSNVGCALSVTGLTVIFQ 219

Search completed: March 21, 2003, 11:38:04
 Job time : 17.4506 secs